2004

Effects of sputter deposition parameters on stress in tantalum films with applications to chemical mechanical planarization of copper

Jeffrey L. Perry

Effects of Sputter Deposition Parameters on Stress in Tantalum Films with Applications to Chemical Mechanical Planarization of Copper

By

Jeffrey L. Perry

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science in Microelectronic Engineering

Approved by

Dr. Michael A. Jackson _____ Date: _____
(Thesis Advisor)

Dr. Santosh K. Kurinec _____ Date: _____
(Committee Member)

Dr. Richard L. Lane _____ Date: _____
(Committee Member)

Dr. Christopher R. Hoople _____ Date: _____
(Committee Member)

Effects of Sputter Deposition Parameters on Stress in Tantalum Films with Applications to Chemical Mechanical Planarization of Copper

Signature of author: _____

Date: _____

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Attempts to introduce a CMP process for copper damascene features at Rochester Institute of Technology were stymied by adhesion failures of the Ta/Cu film stack. This work was undertaken to investigate the effect of stress in the films on adhesion and to develop a viable CMP process for Cu damascene technology.

In depth studies of stress as a function of sputter deposition conditions revealed that stress in Ta layers could vary from -1700 MPa compression to +800 MPa tensile for deposition pressures over a range of 2-20 mTorr for films having a nominal thickness of 0.25 μm. For a fixed pressure, stress could vary from -1500 to +800 MPa for thicknesses ranging from 24 to 225 nm. More importantly, target aging was shown to result in a change in stress for fixed deposition parameters, such as pressure and power. Control of the stress in these films is critical as a substantial difference in CMP removal rates for tantalum films having -400 to -1200 MPa of compressive stress was observed. In addition, the top copper layer will adhere to Ta films in a specific range of compressive stress. A 50 nm film stack of TaN/Ta with varying thickness ratios of the two metals was fabricated that exhibited nearly constant compressive stress. This deposition process for the TaN/Ta barrier layer was developed utilizing fixed voltage, not power as the deposition parameter. These studies resulted in a sputter process for TaN/Ta/Cu that exhibited good adhesion to SiO_2, both for blanket and patterned films.

A copper damascene process has been developed using a film system that adhered well to SiO_2. Wafers were characterized for planarity both within die and within wafer, as well as wafer-to-wafer. The most promising deposition and polish processes were employed to produce a metal gate metal oxide semiconductor (MOS) capacitor and characterized by measuring the maximum electric field of the gate oxide before it would break down. The planarized damascene features were achieved that exhibited $\leq$ 30 nm of topology as viewed by profilometery and AFM. Results of breakdown studies of MOS capacitors were confounded by particulate effects, but the capacitors produced by CMP were on par with sputtered films patterned by photolithography.

ACKNOWLEDGMENTS

I would like to express my gratitude to Dr. Michael Jackson for taking me on as his graduate student and for his guidance throughout this project. I am grateful for Dr. Santosh Kurinec, Dr. Richard Lane and Dr. Christopher Hoople for being on my thesis committee and being willing to donate time to answer my questions. I acknowledge Dr. Tom Blanton for his generous donation of XRD analysis and expertise. I also appreciate the help of Daniel Brown for writing a program to perform stress calculations. This endeavor saved a significant amount of time. I give my most sincere appreciation to my father for providing help in numerous ways.

CHAPTER 1
INTRODUCTION

While the tremendous advances occurring in shrinking integrated circuit (IC) dimensions have resulted in faster devices, the signal propagation through the interconnects between devices has become a concern. For IC's with minimum feature sizes larger than 0.5 μm, circuit delay is primarily due to the device [1]. Minimum feature sizes less than 0.5 μm result in a device delay that is sufficiently small, such that interconnect delay becomes significant. Figure 1 clearly illustrates this phenomenon.

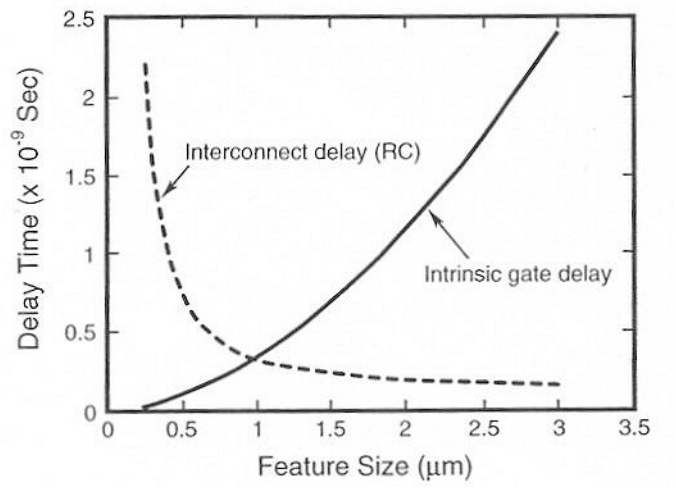


Figure 1: Time delays vs. minimum feature size [2]

1.1 RC DELAY

The interconnect delay in an IC is due to the RC time constant, where R is the resistance of the metal line and C is the capacitance of the interlevel dielectric. The interconnect delay can be estimated using a lumped capacitance model as shown in Equation 1

$$RC \approx \frac{\rho}{t_M}\frac{L^2 \varepsilon_{ILD}}{t_{ILD}} \tag{1}$$

where ρ, t_M, and L are the resistivity, thickness, and length of the interconnect, respectively, ε_{ILD} is the permittivity of the interlayer dielectric (ILD), and t_{ILD} is the thickness of the ILD. From Equation 1 it is apparent that a lower resistivity metal is one option for faster interconnects.

1.2 ELECTROMIGRATION

Copper has emerged as a replacement for aluminum as the interconnect material of choice because of its low resistivity of 1.7 - 2.0 μΩ-cm (~30% less than Al) and its superior resistance to electromigration. Electromigration is the physical movement of conducting metal atoms as a result of momentum transfer from the current-carrying electrons ("electron wind"). This diffusion of interconnect material is faster along grain boundaries, and it can cause a buildup of material in some regions, resulting in hillocks, and depletion in other regions, resulting in voids. As a result shorts and opens in interconnects can develop during the normal operation of a circuit. Figure 2 depicts this process.

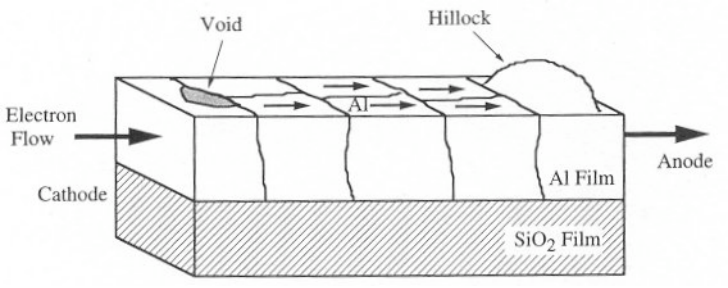


Figure 2: Schematic illustration of electromigration with resultant hillock and void formation in an aluminum line [3]

Electromigration ultimately degrades the reliability of interconnects. This phenomenon occurs faster in aluminum than in copper. The mean time to failure (MTTF) has been modeled empirically with Black's equation

$$MTTF = AJ^{-n} \exp\left(\frac{E_A}{kT}\right) \tag{2}$$

where J is the current density, n is a fitting parameter typically about 2, E_A is the activation energy, T is the absolute temperature, and k is Boltzmann's constant. The

parameter 'A' depends on film structure (i.e. grain size) and processing. Copper has an activation energy of ≈1.25 eV, as compared to aluminum which exhibits a range from 0.5 to 0.8 eV. This larger activation energy translates into reduced failure rates for copper interconnects.

The integration of copper lines into interconnect structures poses several challenges. Copper diffuses quickly into Si and SiO_2 and hence degrades device performance by introducing deep level traps in the semiconductor. In addition, copper has poor adhesion to dielectrics. Therefore, copper interconnects require complete encapsulation by a thin-film layer that functions as both an adhesion promoter and a diffusion barrier.

1.3 TANTALUM BARRIER LAYER

For a given barrier material, there is usually a trade-off between its performance as a diffusion barrier and its performance as an adhesion promoter for copper. If the thin film layer does not react with copper at all, it may exhibit excellent barrier properties but poor adhesion. Interfacial mixing or reaction with copper produces good adhesion but may permit copper diffusion.

Tantalum has been shown to be an excellent candidate as a barrier for copper. It is thermodynamically stable with respect to copper. In fact, Cu and Ta are almost completely immiscible up to their melting points and do not react to produce any compounds [4]. Moreover, the Ta/Si interface has been shown to maintain its stability up to 650^{o}C, with tantalum silicides generally being more stable than their Cu silicide counterparts [5]. When the barrier layer is subjected to high temperatures for a period of time Ta may be susceptible to diffusion barrier failure caused by diffusion of Cu through the Ta, with subsequent formation of Cu_3Si at the Ta/Si interface [4].

1.4 TANTALUM NITRIDE BARRIER LAYER

As in the case of tantalum, tantalum nitrides are thermodynamically stable with respect to Cu. Also, the TaN_x/Si interface is more stable than its Ta/Si counterpart. Most experimental data suggests that the effectiveness of tantalum-based barrier liners increases with higher nitrogen content, at least up to a N-to-Ta ratio of 1:1. This is reported to result from N atoms at interstitial positions that help prevent Cu atoms from diffusing through the TaN layer. In addition, TaN is a good diffusion barrier for F, whereas Ta will react with F. This has implications for copper deposited by MOCVD (metal organic chemical vapor deposition).

Tantalum nitride (TaN_x) encompasses a variety of different phases because it has a defect structure and deviations from stoichiometry are common. In general the structure of tantalum nitrides can be described as close-packed arrangements of Ta atoms with N atoms inserted in interstitial sites. Figure 3 illustrates a tantalum nitride complex and Table 1 lists the various known stoichiometries. Table 2 compares the thermal stability of the various Ta or TaN barriers.

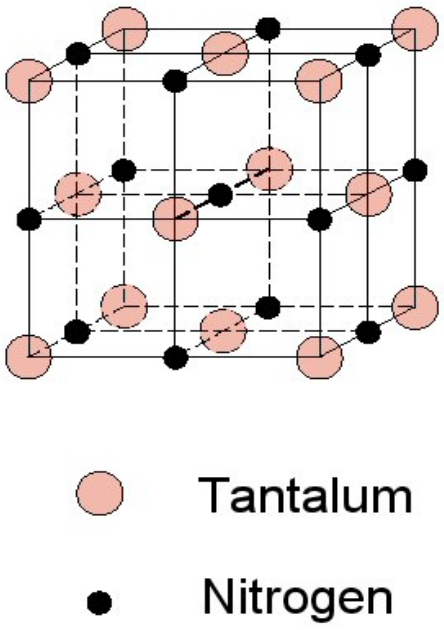


Figure 3: Face-centered cubic tantalum nitride complex [6]

Table 1: Known phases and structures of interstitial nitrides [7]

Tantalum nitride
β-Ta_2N (hcp)
θ-TaN (hcp)
η-TaN (hex)
δ-TaN_{1-x} (fcc)
Ta_5N_6 (hcp)
Ta_4N_5 (hex)
Ta_3N_5 (orth)

fcc = face-centered cubic (close packed)
hcp = hexagonal closed packed
hex = simple hexagonal
orth = orthorhombic

Table 2: Performance of various barrier layers deposited by sputtering [8]

Sample	Barrier Stability
Si/Ta (60 nm)/Cu	600°C, 1hr
Si/Ta (50 nm)/Cu	550°C, 30 min
Si/Ta_2N (50 nm)/Cu	>650°C, 30 min
Si/TaN (100 nm)/Cu	750°C, 1hr
Si/$TiSi_2$ (30 nm)/Ta-Si-N (80 nm)/Cu	900°C, 30 min

1.5 TANTALUM NITRIDE/TANTALUM BARRIER LAYERS FOR IC'S

Current industrial practice utilizes a TaN/β-Ta/Cu approach [9]. The TaN layer ensures adhesion to SiO_2 and the β-Ta layer provides the best adhesion for copper. The β-Ta form also nucleates the copper in an (111) orientation, that is critical for low resistivity and reliability [10]. The presence of N at the SiO_2 surface suggests the possibility of the formation of an oxynitride glass that is more fracture resistant [11].

The Microelectronic Engineering program at Rochester Institute of Technology (RIT) aims to have state-of-the-art processing capabilities and therefore has researched TaN/Ta stacked barrier layers for copper patterned by chemical mechanical planarization (CMP). Previously it has been observed that when a single barrier layer of Ta was deposited by sputtering, adhesion failure would often occur. This provided the

motivation for this work and led to an in-depth study of both the barrier layer(s) and the copper film to develop a reliable process for deposition and planarization of copper. Success in this endeavor would enable future work in multilayer metallization studies, such as dual damascene processing, and replacement gate technology. Being mindful of the benefits of replacement gate technology, it was desirable to ascertain its feasibility at Rochester Institute of Technology using a copper damascene process as an initial step in the investigation.

1.6 REPLACEMENT (DAMASCENE) METAL GATE TECHNOLOGY

In order to improve device performance of sub-100nm CMOS transistors, replacement metal gate technology (RMGT) is being pursued. This technology has the advantage of tighter threshold voltage control and lower gate resistivity than doped polysilicon gates. Moreover, RMGT is not susceptible to polysilicon gate depletion that leads to an increase in the effective oxide thickness and dopant penetration through thin dielectrics. If a damascene process is used to implement this technology, it will have the additional advantage of low temperature processing. This is helpful since very high-k dielectrics become unstable or leaky after high temperature processing (>850^{o}C) [20]. In addition, plasma damage on gate structures will be suppressed which can be generated during reactive ion etching.

CHAPTER 2

BACKGROUND

Previously at Rochester Institute of Technology, it had been observed that when a copper film was deposited on a Ta barrier layer, adhesion failure would often occur. A hypothesis was made that stress may play a role in this adhesion failure, therefore, this work was undertaken to understand and control the stress. Our study utilized PVD for both the barrier layer(s) and the copper film, while in industry copper is often electroplated. Because this study was concerned about the adhesion failure issues, the method of copper deposition was deemed to be of secondary importance. An overview of sputtering fundamentals follows to give insight to the relationship between film stress, adhesion and the process variables that control them.

2.1 FUNDAMENTALS OF SPUTTERING

Sputtering is a vacuum based, physical process, in which the positive ions in a glow discharge strike a target and eject atoms from it by momentum transfer. These atoms travel and are deposited on an adjacent substrate material. Sputtering occurs when the kinetic energy of the incoming ions exceeds the binding energy of the surface atoms of the target.

A simplified DC (diode) sputter system is shown in Figure 4. The target is a disc of the material that is the source of the film to be deposited, or the material from which the film is synthesized (reactive sputtering). Because the target is biased negatively it is also known as the cathode. Typically, up to several kilovolts can be applied to it. The substrate that faces the cathode may be grounded, electrically floating, biased positively or negatively, heated, cooled or some combination of these. It will be biased positively with respect to the cathode, and is referred to as the anode.

After evacuation of the chamber, a gas, usually argon, is introduced and serves as the medium in which a glow discharge is initiated and sustained. In IC production, deposition pressures typically range from 1 to 5 mTorr. After the glow discharge is

sustained between the electrodes, the ion current component will result in the sputtering of the target, and the film condenses on the substrate.

Magnetron-based sputter tools deposit thin films at much higher rates than simple diode systems. A DC magnetron is basically a magnetically enhanced diode in which the spatial relationship of electric ($\mathbb{E}$) and magnetic ($\mathbb{B}$) fields is engineered to confine secondary electrons produced by Ar^+ bombardment of the target. Restricting these electrons close to the target surface increases their probability of ionizing the Ar working gas, which in turn results in a more intense plasma discharge that can be sustained at low pressures. In addition, they operate at lower pressures where scattering and impurities from the gas phase are minimal. A more thorough overview of sputtering may be found in Reference [38].

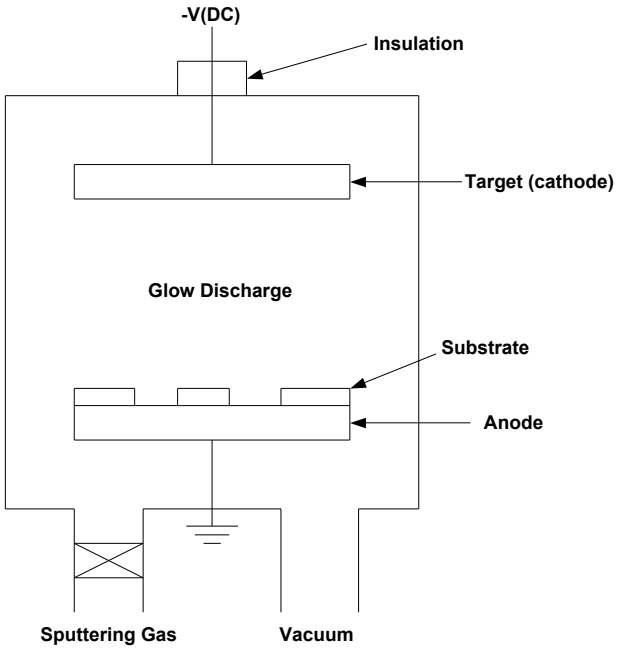


Figure 4: Schematic of a simplified DC sputter system

2.2 REACTIVE SPUTTERING

Reactive sputtering is the method in which thin films of compounds are deposited on substrates by sputtering from metallic targets using both the inert working gas and a reactive gas. The dynamics of this process as a function of reactive gas flow is depicted in Figure 5. At low flow rates all of the reactive gas is incorporated into the deposited film (metallic mode). As the gas flow rate is increased, a threshold is reached where the target surface experiences compound formation. When this compound formation exceeds the removal rate of material, a threshold is reached that is accompanied by a sharp decrease in sputtering rate and discharge voltage (compound mode). This decrease is due to the fact that compounds generally have lower sputtering rates and higher secondary-electron emissions. Additionally, a third cause for the drop in sputtering rate is due to less efficient sputtering by reactive gas ions than by inert ions. When the reactive gas flow is sufficiently reduced the system will revert back to metallic mode. However, metallic mode sputtering will not commence at the same flow rate, because the compounds remaining on the target need to be removed before normal sputtering can resume.

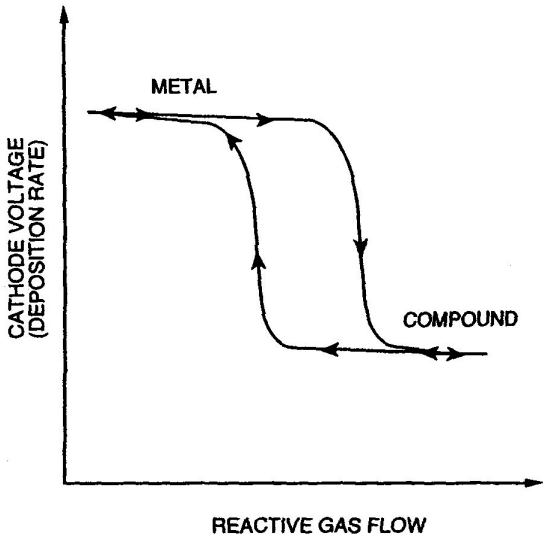


Figure 5: Generic hysteresis curve for voltage vs. reactive gas flow rate [12]

2.3 COPPER DAMASCENE PROCESS

There are no working processes that allow copper to be plasma etched resulting in anisotropic sidewalls (which is necessary for submicron features). This is because copper cannot form a volatile compound in the gas phase at low temperatures, without severe contamination of the devices. Therefore, the most effective method of pattering copper is to etch the ILD first followed by deposition of a liner and copper. Chemical Mechanical Planarization (CMP) is then used to remove all the metal except which that which is imbedded in ILD traces. This process is commonly referred to as the copper-damascene process and is illustrated in Figure 6.

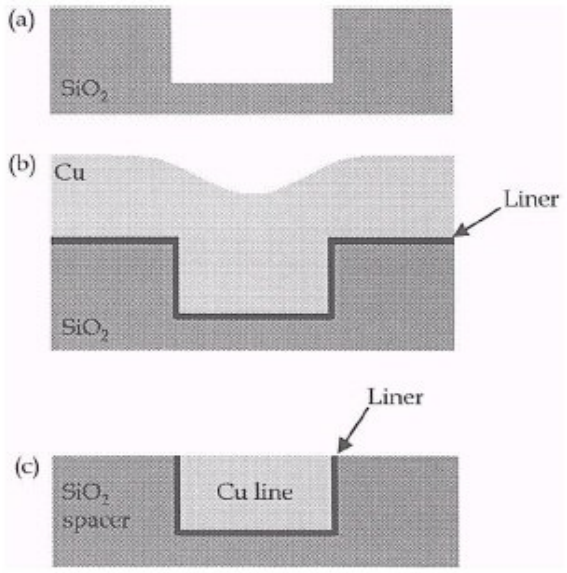


Figure 6: Cross-sectional view of the different process steps in the copper damascene process: (a) Patterns are etched in the ILD, (b) Liner and copper are deposited, (c) CMP removes the copper overburden and liner [15]

2.3.1 CHEMICAL MECHANICAL PLANARIZATION (CMP)

A schematic diagram of the Chemical Mechanical Planarization (CMP) operation is shown in Figure 7. The wafer is pressed against a polishing pad supported by a plate. Both the wafer and platen are rotated. A polishing slurry, consisting of submicron-sized particles suspended in an aqueous medium and other chemical agents, is dispensed onto the pad. Centrifugal force distributes the slurry across the pad forming a thin sheet of liquid and saturating the pad. The abrasives in the slurry mechanically abrade the chemically modified surface layers, resulting in material removal.

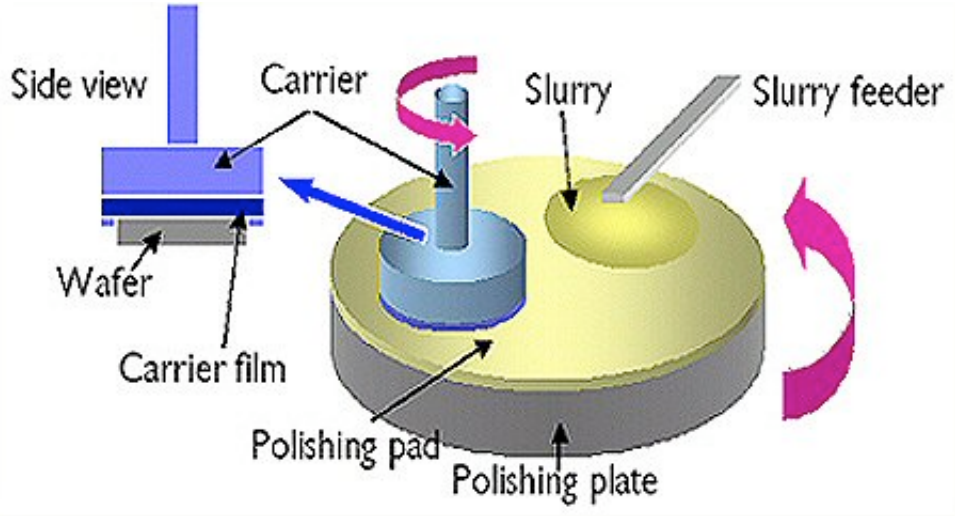


Figure 7: Schematic representation of typical CMP process [13]

The most frequently referenced expression for modeling the removal rate from CMP is the Preston equation

$$R = KPV \tag{3}$$

where R denotes the polish rate in m/s, P is the applied downward pressure in Pa, V is the linear velocity of the wafer relative to the polishing pad in m/s, and K is a proportionality constant that depends on the tool, chemistry, and temperature.

2.3.2 MODIFIED PRESTON EQUATION FOR COPPER CMP

Figure 8 illustrates that the Preston equation is not applicable when polishing copper. The removal rate is shown to have a non-zero intercept for both zero revolutions./min and/or zero down force. To take this into account Luo, et al. [14] fitted the Preston equation with two additional parameters as shown in Equation 4

$$R = (KP + B)V + R_C \tag{4}$$

where R_C is a constant to represent the non-zero polish rate measured at zero downward pressure and/or zero rotational speed, and B is an additional term proportional to the rotational speed. The value of the constant R_C reflects the purely chemical reactivity of a copper slurry. This is necessary because copper does not form a passivating layer in an acidic medium which is the chemical state of most copper slurries.

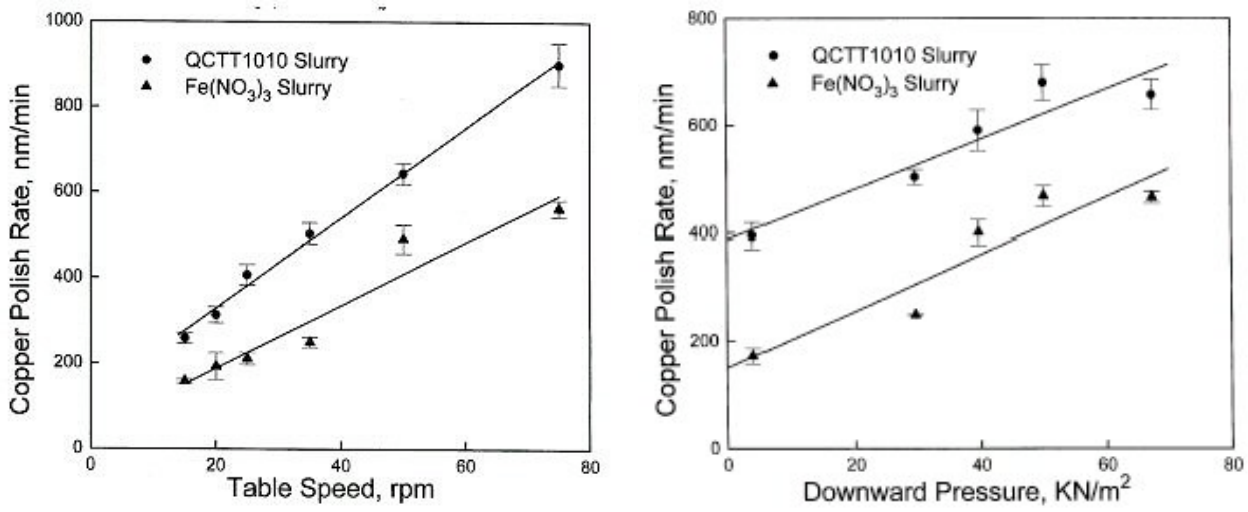


Figure 8: Effect of rotational speed and pressure on copper polish rate [14]

2.4 STRESS IN THIN FILMS

The stress in films is an increasingly important technological issue from the standpoint of reliability and performance in IC processing. Large stress levels may give rise to void formation upon subsequent thermal cycling where the strain state provides a driving force for grain boundary diffusion. Variations in stress levels between films resulting from changes in the sputtering process could influence removal rates in CMP. Finally, if the stress in a film is too large, the film may peel from the wafer surface.

Stress may be either compressive or tensile, as shown in Figure 9. When a film is trying to expand on a substrate it will be in a state of compression and have a negative radius of curvature. Conversely, if the film is trying to contract it will be tensile in nature and have a positive radius of curvature.

No Stress

Infinite Radius of Curvature

Tensile Stress

Positive Radius of Curvature

Film: Trying to Contract

Compressive Stress

Negative Radius of Curvature

Film: Trying to Expand

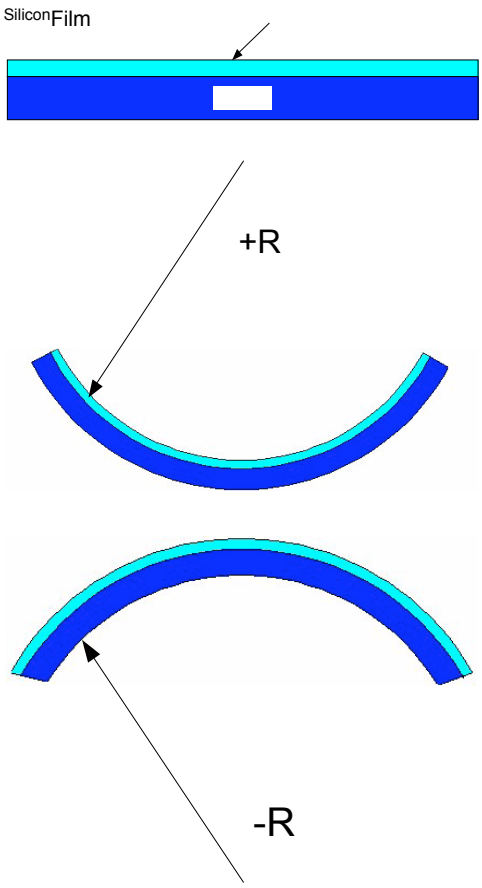


Figure 9: Two types of stresses in thin films

Virtually all sputtered coatings are in a state of stress. The total film stress is composed of extrinsic stress, thermal stress and intrinsic stress.

$$\_\_total = \_\_extrinsic + \_\_thermal + \_\_intrinsic \quad (5)$$

Extrinsic stress is induced by external factors and results from interactions between the deposited material and the environment which are subsequent to deposition (i.e., this type of stress may arise from adsorption of water vapor in porous films exposed to room air immediately after removing the samples from the deposition chamber) [39]. The thermal stress is due to the difference in the thermal expansion coefficients of the coating and substrate materials. The intrinsic stress is due to the accumulating effect of the crystallographic flaws that are built into the film itself during deposition. In hard, high melting point materials such as tantalum, when deposited at low temperatures the intrinsic stress tends to dominate over thermal stresses. Conversely, in soft low melting point materials such as aluminum, bulk diffusion tends to relax the internal stresses and prevent their accumulation. Therefore, thermal stress is dominant. These trends are depicted in Figure 10.

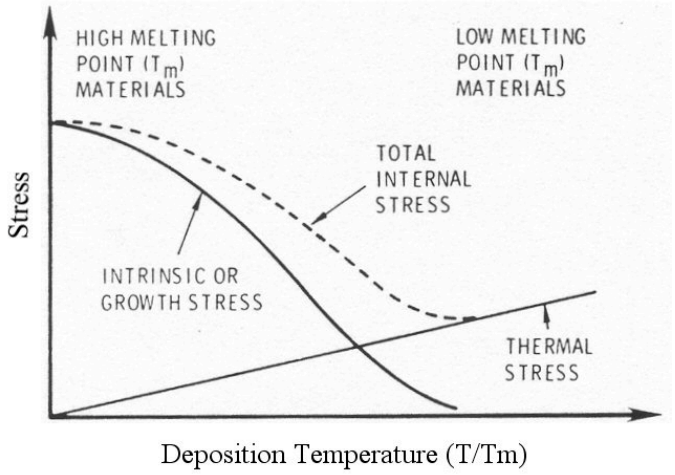


Figure 10: Trade-off between thermal and intrinsic stress in thin films for high and low melting point materials [16]

2.4.1 THERMAL STRESS

When a coated substrate is at a temperature that is different from its temperature during deposition, a thermal stress will be present as a result of the differences in the film and substrate thermal expansion coefficients. For a one-dimensional approximation, the thermal stress is given as

$$\sigma_{Thermal} = E_f (\alpha_f - \alpha_s)(T_s - T_A) \tag{6}$$

where E_f is Young's modulus in Pa, α_f and α_s are the coefficients of thermal expansion for the film and substrate, respectively, in $1/^oC$, T_s is the substrate temperature during deposition, and T_A is the temperature during measurement. A positive value of $\sigma_{thermal}$ corresponds to tensile stress where a negative one corresponds to compressive stress.

2.4.2 INTRINSIC STRESS

Intrinsic stress is less well understood. It can be defined as the component of the total measured stress that cannot be attributed to external or thermal stress. The intrinsic stress can depend on such variables as the substrate temperature, deposition rate, film thickness, and background chamber ambient. Intrinsic stress develops in films during the growth process. The magnitude of intrinsic stress is related to the microstructure of the film which, in turn, depends on the kinetic energy of atoms condensing on the substrate or other energetic species impinging on the surface during film growth.

Figure 11 shows a standard zone diagram (SZD) for sputtered metallic coatings. The diagram is based on experimental data and shows the effect of argon pressure and substrate temperature on the morphology of metallic coatings. Notice that the substrate temperature variable is really the ratio of the substrate temperature to the melting point of the bulk material (T/T_M). This ratio is referred to as the homologous temperature.

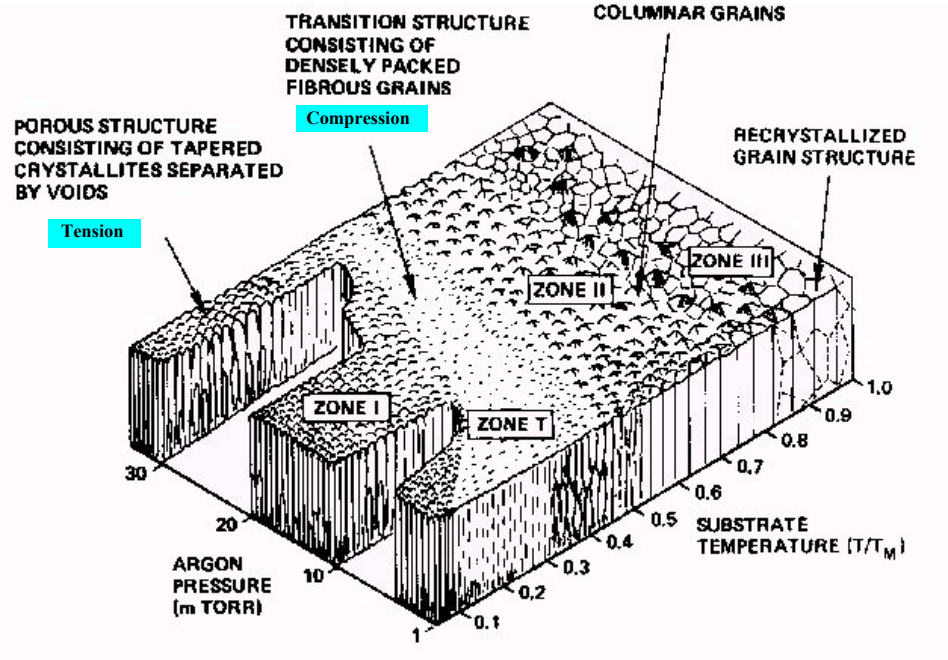


Figure 11: Schematic representation of the influence of substrate temperature and argon pressure on the microstructure of metal coatings deposited by cylindrical magnetron sputtering [12]

There are four different morphology regions in the SZD. Below each of the four regions are listed with a brief description.

Zone T: Dense array of poorly defined fibrous grains.
Zone 1: Well-defined columnar structure (one-dimensional crystals) with voids in between.
Zone 2: Narrowing of void structure due to increased surface mobility of adatoms.
Zone 3: Recrystallization of materials due to bulk diffusion of adatoms within the film.

In IC processing only Zones T and 1 are of practical importance for metallization. Coatings in Zone 1 are characterized by a structure consisting of tapered crystals separated by open, void boundaries. This structure results from shadowing because high points on the growing surface receive a larger coating flux than valleys, particularly when a significant oblique component is present in the arriving coating flux. This is enhanced by elevated working gas pressures. Zone 1 films are often characterized by tensile stresses. Coatings in Zone T have a dense fibrous structure with a smooth, highly

reflective surface. They form when the coating flux arrives in a direction that is largely normal to the substrate surface so that shadowing is minimized. Zone T films are mostly characterized by compressive stresses.

2.5 MODELS FOR THIN FILM STRESS

The Grain Boundary Relaxation (GBR) model, is the one most often used to explain the tensile stress in polycrystalline films. The model is based on the following physical argument; as the film growth progresses through morphological stages (from isolated atomic clusters to a continuous film), interatomic attractive forces acting across the gaps between contiguous grains cause elastic deformation of the grain walls [17]. This is illustrated in Figure 12. The deformation is counterbalanced by the intragrain or intracolumn tensile forces imposed by the constraint caused by the adhesion of the film to the substrate surface. Implicit in the model is the assumption that the adhesive forces to the substrate exceed the intergrain attractive forces.

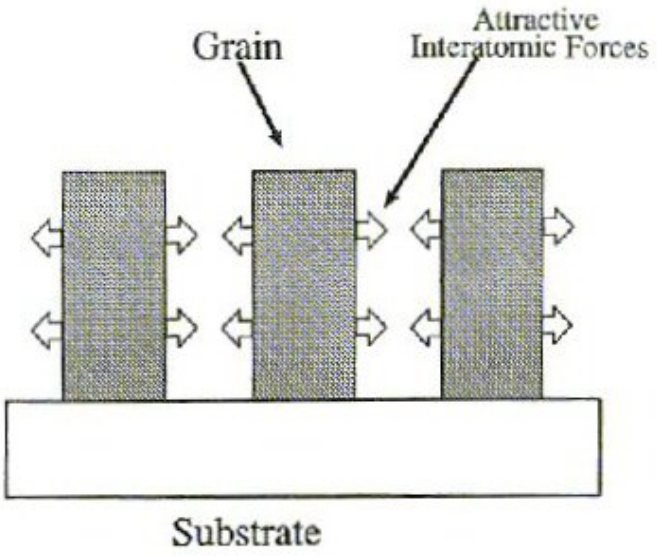


Figure 12: Schematic representation of the grain boundary relaxation model [18]

Two models have been used extensively to account for compressive stress in sputtered films. The first uses the atomic peening mechanism and the second involves

film impurities. Atomic peening occurs at low pressures where sputtered atoms and reflected argon atoms impinge on the film at near normal incidence and with high energies. This is because at lower pressures there are fewer collisions within the plasma. In striking the coating with high momentum, incident atoms drive the surface atoms of the film closer together, imbedding themselves in the film. The end result is compressive stress. An exact mechanism for film impurities has not been identified. However, the model is based on the concept of lattice distortion produced by one of the following: (1) incorporation of atoms of a different size from the film, (2) reaction at grain boundaries that produce a phase with a different molar volume, or (3) grain surface energy reduction. Presently, there are no quantitative impurity models. Stress data is usually correlated spectroscopically with the impurity concentration [17].

Studies report that oxygen impurities in sputtered films produce compressive stress [17]. Water on the other hand, can produce either tensile or compressive stress. The stress associated with water vapor is dependant on the sign and magnitude of the dipole-induced stress between water and the film. This dipole in turn depends on the microstructure and the chemical state of the absorbing surface. For an inert gas there is no correlation between stress and gas incorporation. Figure 13 shows no appreciable difference in the strain of sputtered films for a 2-decade variation in argon content for a variety of materials. The relationship between stress and strain is: stress = (strain) x (Young's modulus). This assumes an elastic deformation of the material.

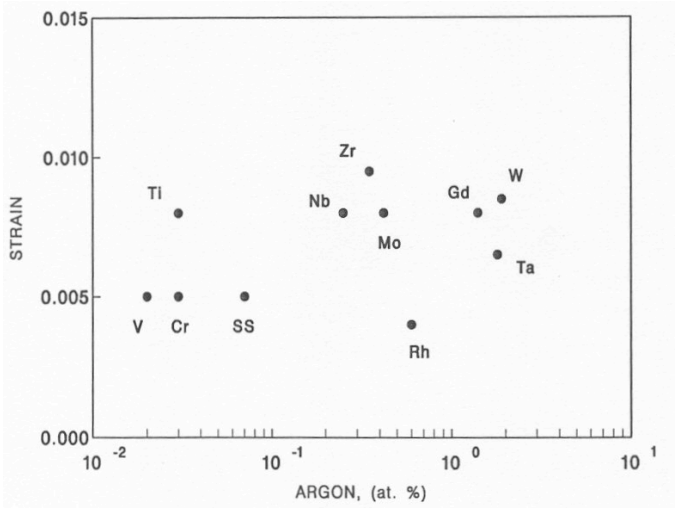


Figure 13: Compressive strain vs. the argon content of sputtered films [17]

2.6 EFFECT OF DEPOSITON PRESSURE ON STRESS

Numerous studies have been conducted demonstrating the function of pressure on stress [17]. A typical stress-pressure curve for sputtered films is presented in Figure 14. There are three characteristic regions commonly observed. As the plasma pressure is reduced at high pressures, the tensile stress will increase as a result of both resputtered oxygen based contaminants (which tend to produce compressive stress) and microporosity annihilation occurring simultaneously. This is the region labeled (1) in Figure 14. Region (2) is characterized by an abrupt transition from tensile to compressive stress, followed by a highly compressive stress region (3) that forms at low pressures. The sharp transition in the second region is caused by the atomic peening mechanism. At lower pressures the target atoms can more easily reach the substrate with fewer collisions within the plasma. Therefore, they have greater momentum and bombard the substrate at near normal incidence [17, 18]. The end result is film densification and increased compressive stress.

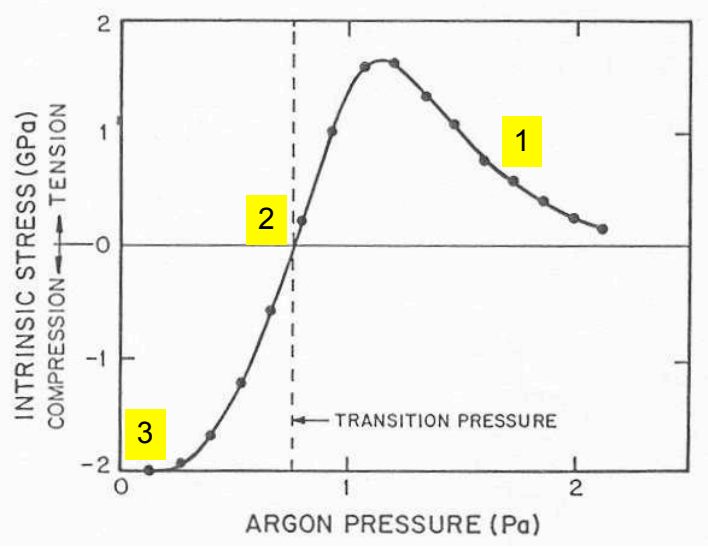


Figure 14: Typical stress-pressure curve for a metal film at low T/T_m deposition temperatures [17]

In a sputter deposition process almost every variable exhibits a stress reversal at some point in its range. Table 3 lists several of these variables. The gas pressure and cathode power are among the easiest to control and allow the experimenter to "engineer" desired states of film stress.

Table 3: Effect of process variables on stress for low T/T_M deposition [19]

"Compressive"	Variable	"Tensile"
Negative	Substrate bias	Positive
Low	Gas Pressure	High
Low	Gas atomic mass	High
High	Target atomic mass	Low
Normal	Angle of deposition	Oblique
Oblique	Angle of emission	Normal
Cylindrical	Target shape	Planar
High	Cathode Power	Low
	Magnetron parameter	
	Substrate proximity	
	Substrate motion	
	Reactive contamination	

It is important in sputtering to know at which pressure at which the stress reversal occurs for a given process. This is illustrated in Figure 15. The data indicates that certain materials, such as tantalum, tend to produce only compressive films while other materials, such as aluminum, form tensile films.

2.7 FILM GROWTH SUMMARY

Figure 16 summarizes much of this discussion. At low working gas pressure, the argon and the sputtered atoms easily reach the substrate due to fewer collisions within the plasma (long mean free path length). Thus, they retain greater momentum and bombard the substrate at near normal incident. In this situation, atomic peening readily occurs resulting in films exhibiting compressive stress. At higher gas pressures, the argon and target atoms bombard the substrate with less momentum and at more oblique angles because of increased collisions in the plasma. This reduces adatom mobility during film growth resulting in atomic shadowing and creation of tensile stress due to grain boundary

relaxation. In general, a slightly compressive films are desired. This is because these films tend to have better conductive, adhesive, and reflective properties.

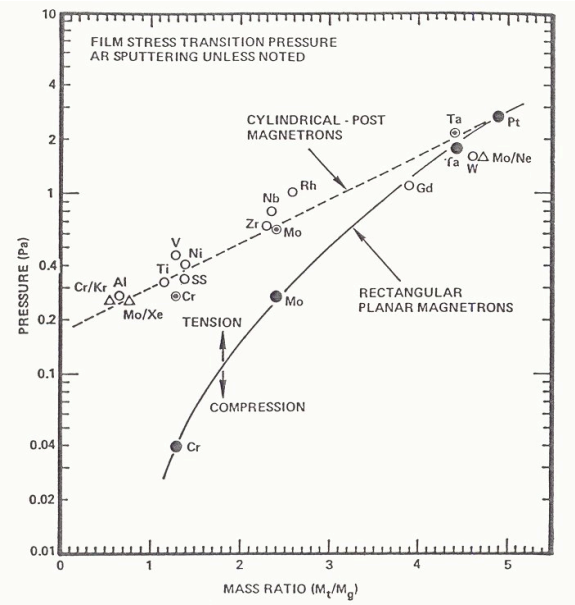


Figure 15: Stress transition pressure for different materials [17]

To successfully fabricate a TaN/Ta/Cu film stack for use in IC manufacturing, careful control of the film stress and an understanding of its effects on adhesion and CMP performance was desired. To achieve that end, numerous measurement or characterization techniques were employed. The remainder of this chapter consists of a brief overview of these techniques to provide enough background for the Experimental and Results Chapters.

2.8 STRESS MEASUREMENT

A common method for determining the stress in a film is by quantifying the change in the radius of curvature for a substrate after a film is deposited. Assuming the film thickness is much smaller than the thickness of the substrate and that the film is in an isotropic, biaxial stress state ($\sigma_z = 0$, $\sigma_x = \sigma_y$) then the biaxial stress in the film, σ_f, can be approximated by

$$\sigma_f = \frac{1}{6R} \frac{Y_s x_s^2}{(1-\nu_s) x_f} \tag{6}$$

where R is the radius of curvature in m, Y_s is Young's modulus of the substrate, ν_s is Poisson's ratio for the substrate, x_s is the substrate thickness, and x_f is the film thickness. In practice, the curvature is measured before and after the film is deposited resulting in 1/R being determined by $(1/R_{final} - 1/R_{initial})$. For (100) silicon $Y_s/(1-\nu_s)$ = 181 GPa.

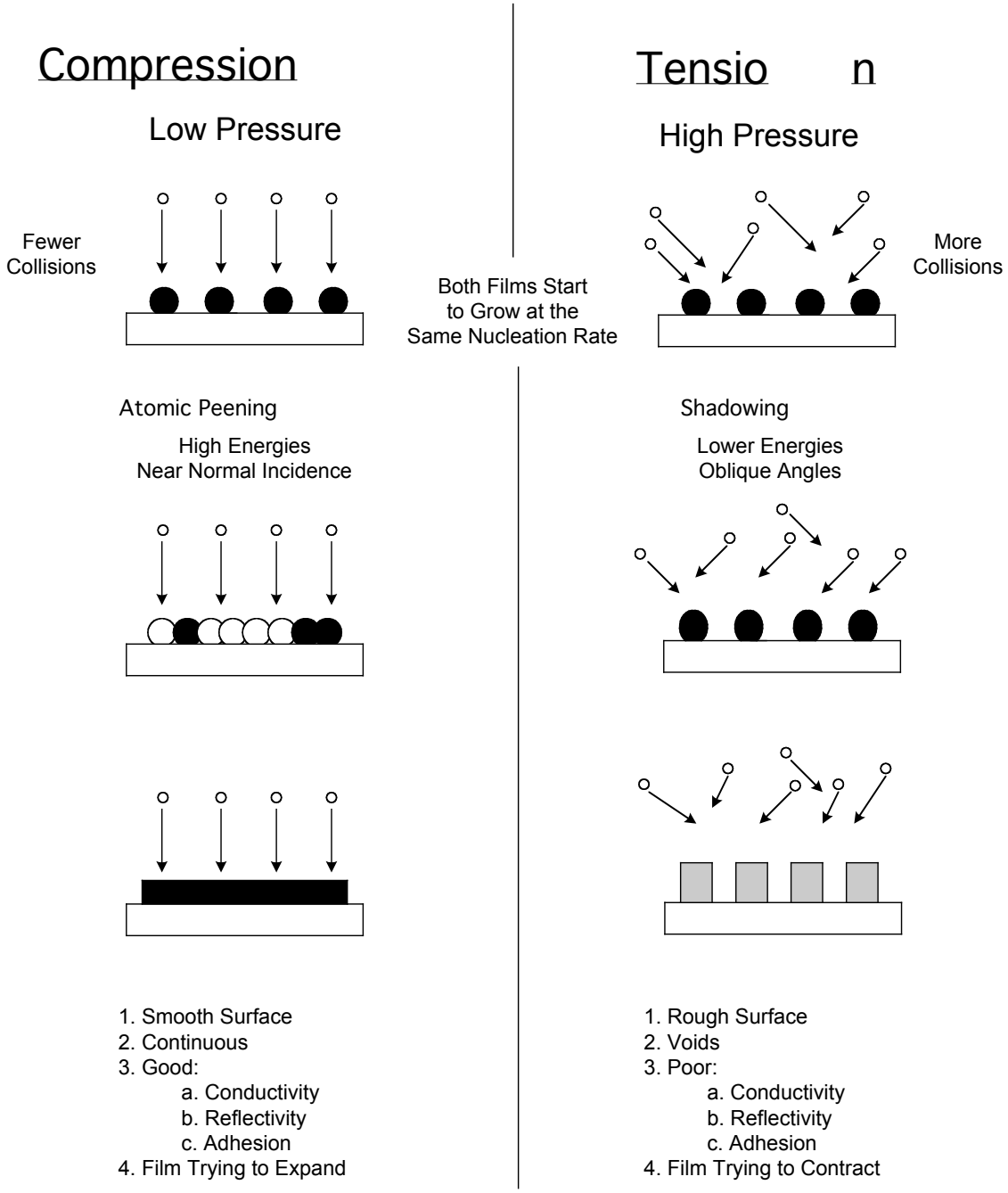


Figure 16: Basic idea of film growth at low and higher working gas pressures

2.9 X-RAY DIFFRACTION

X-ray diffraction (XRD) techniques are well developed for the characterization of crystallinity in materials such as metals, ceramics, polymers and other inorganic and organic compounds. XRD can be used to identify the phases present in a sample and provide information on the physical state of the sample such as grain size and texture. Most XRD techniques are rapid and nondestructive.

When a monochromatic x-ray beam is incident on the surface of a crystal, it is diffracted and constructive interference is observed when the angle of incidence has certain values. These angles depend on the wavelength and the lattice constants of the crystal. A simplified diagram is given in Figure 17 where a crystal is represented by a set of parallel planes corresponding to the atomic planes. The incident beam is diffracted partially at each of these planes, and then collected simultaneously at a distant detector. The interference is constructive only if the difference between the paths of any two consecutive rays is an integral multiple of the wavelength. That is, path difference = $n\lambda$, where λ is the wavelength and n is a positive integer. XRD obeys Bragg's law which is

$$2d \sin\theta = n\lambda \tag{8}$$

where d is the interplanar distance and θ the glancing angle. By measuring θ and λ, it is possible to determine d and eventually the crystal structure.

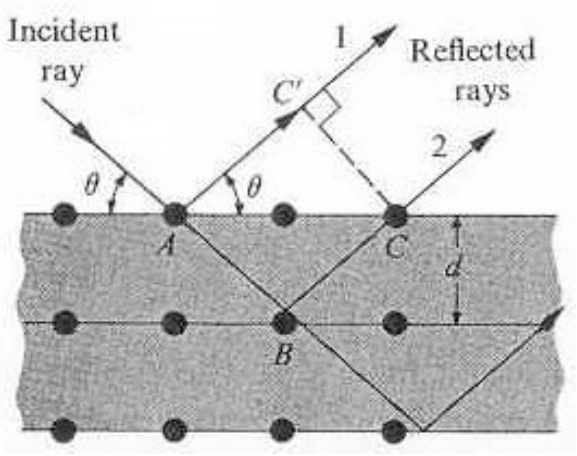


Figure 17: Reflection of x-rays from a crystal [22]

2.10 ATOMIC FORCE MICROSCOPE

After the CMP process, the surface flatness must be determined. The Atomic Force Microscope (AFM) is an instrument capable of visualizing surface features on a molecular scale. It operates on a very similar principle to that of a profilometer. A probe tip is suspended from a cantilever whose deflection is then used to monitor the surface forces along a scan.

Figure 18 is a schematic diagram of an AFM. A probe tip is mounted onto a cantilever beam that is flexible enough to allow the microscopic stylus to respond to the variations of the sample profile. The cantilever is rigid enough, however, to restore contact between the probe and the sample if they become separated. The most common detection scheme is the use of an optical lever. Here light is reflected from the back of the cantilever and its deflection due to the displacement of the cantilever is monitored with a photodetector. With an optical AFM, forces from 10^6-10^9 Newtons can be routinely measured. With raster scanning of an AFM, surface topology is visualized. This allows for the measurement of average roughness, particle size and the calculation of distances and angles between objects.

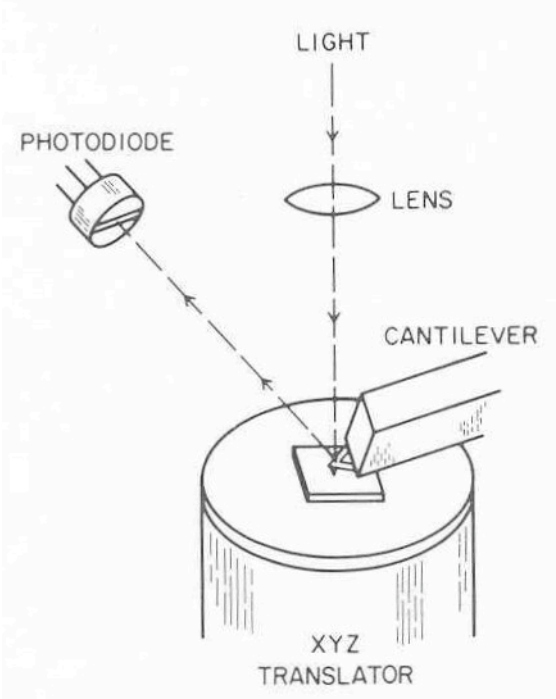


Figure 18: Schematic diagram of an AFM with an optical lever [21]

2.11 SURFACE CHARGE ANALYZER

Before the metal films are integrated into a MOS device, the quality of the gate oxide should be determined. The Surface Charge Analyzer (SCA) is an electro-optical method that allows for the rapid and nondestructive characterization of the electronic properties of a bare semiconductor surface or one covered with an insulator. The SCA technique may be treated as an electro-optical equivalent of the metal-oxide-semiconductor capacitance-voltage (MOS C-V) technique. However, unlike the MOS C-V technique, it does not require the presence of an oxide or the preparation of special test structures resulting in the elimination of processing steps and faster turn around times.

In the SCA method, the electronic properties of the surface are determined from measurements of the alternating surface charge generated as a function of an electric field capacitively applied to the semiconductor-insulator structure. The schematic of the SCA measurement technique and the various charges at the semiconductor surface are depicted

in Figure 19. The ac signal is generated with a beam of pulsed light, incident on the wafer surface, with a photon energy greater than the semiconductor bandgap. The semiconductor surface is swept from accumulation through depletion into inversion (or vice versa) with the superimposed electric field. The illumination is adjusted to a level at which the measured ac signal is proportional to the incident light intensity. Under such conditions, the induced signal is proportional to the depletion layer width. By measuring the depletion layer width dependence on induced charge, while the surface is externally biased, the surface doping concentration, the oxide charge, and the density and energy distribution of the interface states can be determined. A detailed description of the theory of operation can be found in reference [23].

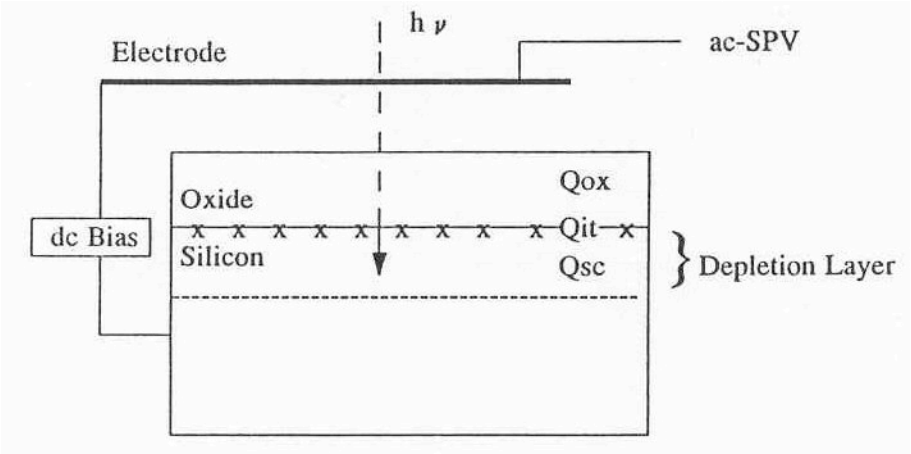


Figure 19: Schematic of SCA measurement technique. Q_{ox} = oxide charge, Q_{it} = interface trap charge, Q_{sc} = semiconductor space charge [23]

2.12 MOS CAPACITOR

The most dominant electrical technique to characterize the performance of MOS gates is the Capacitance-Voltage or CV method. This technique is very widely used and provides a large amount of information about dielectric films and the interfaces that such films make with underlying semiconductors. The basic structure used to make CV

measurements is the MOS capacitor. It consists of a semiconductor, a dielectric layer and a conducting electrode.

The basic measurement is illustrated in Figure 20. Here we will consider the specific case of a metal/SiO_2/n-type silicon structure with only a small number of charges or defects. In this structure the metal part of the structure is called the gate and is biased with a DC voltage with respect to the substrate.

Consider the first step called accumulation with a gate voltage of $+V_G$ (Fig. 20a). Since the substrate is n-type, the positive gate voltage will attract the majority electron carriers to the silicon surface. Here the capacitance is simply the oxide capacitance, C_{ox}.

Next in depletion, a negative gate voltage is applied (Fig. 20b). This negative voltage repels the majority electron carriers from the surface. This creates a depleted region with only positively charged donor atoms. The depleted region has a width x_D and sets up another capacitance called the depletion capacitance, C_D. Now the overall capacitance is C_{ox} in series with varying capacitance C_D. The C_D voltage varies with the size of the depletion width, this in turn is controlled by the gate voltage.

In inversion a larger negative gate voltages is applied. This causes the silicon surface to invert from n-type to p-type (Fig. 20c). The negative voltage on the gate attracts the minority hole carriers in the substrate to the surface. When enough of them are present, they will form an inverted layer of p-type carriers. The gate voltage at which this occurs is called the threshold voltage. At this point, the depletion width x_D stops expanding and reaches a maximum value of x_{DMax}. This causes the series capacitance to reach a minimum (assumes the curve is generated at high frequency).

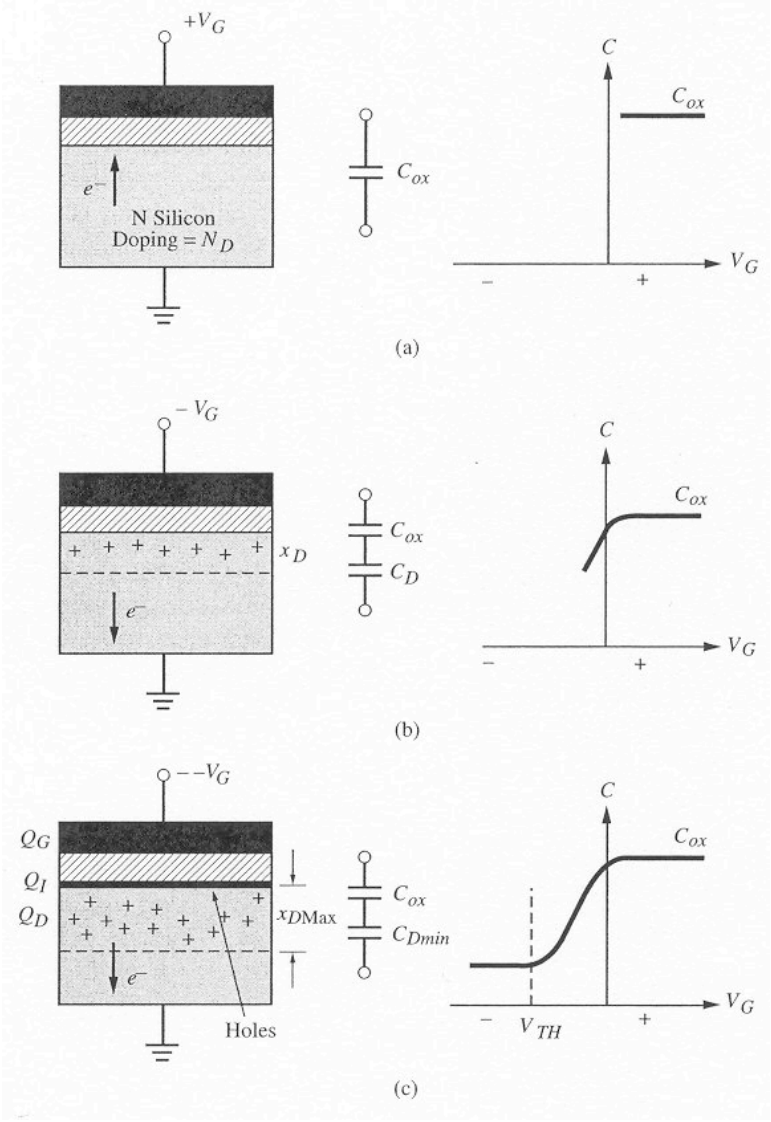


Figure 20: MOS capacitor structure and resulting CV plot for (a) accumulation, (b) depletion, and (c) inversion [3].

With the aforementioned knowledge in film stress, sputter and CMP technologies and various measurement techniques the background is set for further experimental study. It is now possible to investigate the adhesion failure issue of a Ta barrier layer when copper is deposited onto it.

CHAPTER 3

EXPERIMENTAL DETAILS

This work was completed in two phases. First an in-depth study of stress in Ta, TaN, and Cu films was conducted with the intent of identifying the effects of deposition conditions upon film adhesion. Second, a CMP process was developed to pattern the film stack, and that process was then employed to fabricate MOS capacitors as a means of determining the feasibility of a replacement gate technology. The details of each experimental process now follow.

3.1 SPUTTERING

DC magnetron sputtering was performed in a CVC 601 sputter tool to deposit all metallic films. Tantalum was sputtered using a 4-inch target while an 8-inch target was used for copper. In the system, the targets sputter upwards toward the substrate holder, which is being rotated at 14.5 rev/min. The target-to-substrate distance was 5.5 cm. A cold cathode ion gauge and a baratron (capacitance diaphragm) were used to the measure the base and working pressures, respectively. The system used an ENI RPG-50 power supply. The base pressure was $\leq$ 1-2x10^{-5} mTorr and was achieved with a cryopump.

The films were deposited onto 4-inch silicon (100) wafers that had a 1 μm thick, steam oxide that had been grown in a Bruce furnace. Before every deposition the wafers were put through a dilute HF dip (100:1, H_2O:HF). This process step cleaned the surface and allowed for better adhesion. This is because the oxide wafers were often in storage for extended periods of time prior to use, and the short etch was used to produce a "fresh" surface for the deposition. The end result was better adhesion.

Before each target was used to deposit a film, it was conditioned to eliminate unwanted oxides or other contamination from its surface. Moreover, a 40-minute bake

out step at 250^{o}C was conducted for each run to remove adsorbed water vapor from the substrate surfaces that could be detrimental to film adhesion. Specific times and settings are reported in Chapter 4.

3.2 FILM CHARACTERIZATION

3.2.1 Thickness

Film thickness measurements were made using a Tencor P2 Profilometer. Step heights were created by marking a line on the wafers prior to deposition with a Sharpie pen. Afterwards, the deposited metal was removed using isopropyl alcohol and a Q-tip. To obtain consistent step height measurements the wafers were orientated in an identical manner on the platen and the film thickness was taken across the same diameter for each deposition. Figures 21 and 22 show the positioning of the substrate on the platen and the site pattern for the thickness measurements, respectively. This approach was chosen to obtain accurate thickness uniformity. If the thickness readings were taken parallel to the platen rotation, misleadingly "good" uniformity profiles would be obtained because the rotation of the substrate averages out the thickness variations in this direction. To prevent this they were taken perpendicular to the platen rotation to get a more realistic or "worst case" thickness profile.

Figure 22 shows nine measurement points separated by 1 cm, with 1 cm of edge exclusion. Thickness data was taken as the average of points X3-X7. Only the middle 40 mm was used to compute the average because the film thickness would tend to thin out near the edges of the wafer. Moreover, it was found that there was no statistical difference in film uniformity for various orientations of diameter if the middle portion was used for film thickness. The decision to use the middle portion was correlated with the fact that stress measurements were carried out with a 50-mm wafer scan in the center of the wafer. This helped minimize stress variations that are dependent mainly on the wafer bow caused by the deposited film in the most central portion of the wafer.

3.2.2 Stress

The P2 determined the film stress by measuring the change in the radius of curvature of the silicon wafers before and after deposition. Figure 23 shows how six different stress readings at various angles were taken. The average of these values was recorded as the film stress. Wafer thickness, which is necessary for accurate determination of stress, was measured with a micrometer. The P2's accuracy was ± 17 MPa.

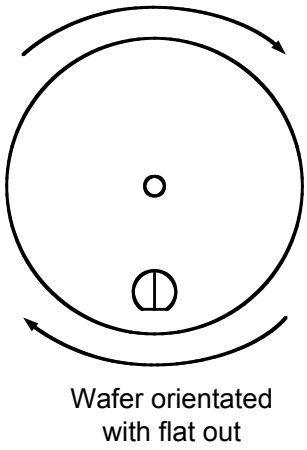


Figure 21: Wafer orientation during sputtering to obtain consistent thickness measurements

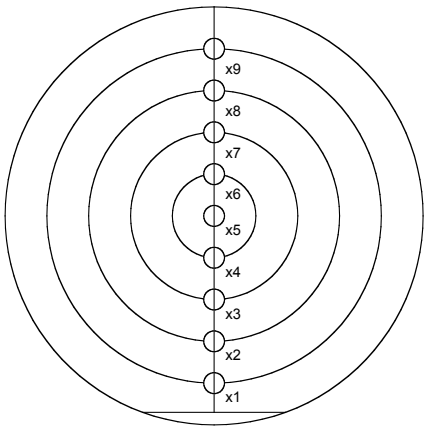


Figure 22: Schematic of where film thickness measurements were taken (perpendicular to wafer flat)

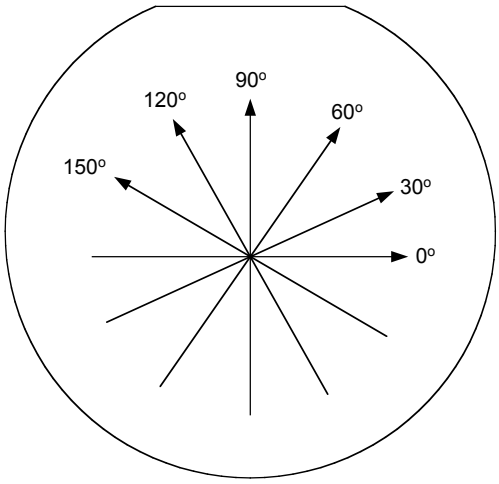


Figure 23: Diagram of stress measurements

3.2.3 XRD and resistivity

Samples were sent out for X-ray Diffraction Analysis (XRD) analysis. They were analyzed with a Rigaku D2000 diffractometer equipped with a copper rotating anode, diffracted beam graphite monochromator tuned to CuKα radiation, and a scintillation detector. Film resistivity measurements were made using a 4-point probe and a CDE ResMap 178.

3.3 CHEMICAL MECHANICAL PLANARIZATION

Polishing of films was done with a Strausbaugh CMP tool. The pad used was a Rodel stacked pad with a perforated IC1000 on top of a SUBAIV. EKC copper slurries were used. There were two phases of slurry.

- Phase I slurry consisted of EKC MicroPlanar CMP9001 and CMP9007. CMP 9001 is the abrasive consisting of alumina particles and CMP9007 is a copper oxidizer solution. It is designed to have a high selectivity for copper and stop on a tantalum barrier layer.

- Phase II slurry is made from MicroPlanar CMP 9003 and CMP 9011. CMP 9003 is a silica abrasive while CMP 9011 is the oxidizer solution. It is made solely to remove the barrier layer. It has virtually no removal rate for copper.

The CMP conditions are listed in Table 4 (if necessary, major deviations from this are appropriately indicated). There was no in-situ rinsing of the wafers on the CMP tool. This was performed manually after polishing. In addition, the quill speed was set close to the table rotation speed but not identical. This was done to prevent the creation of a resonance during polishing between the polish table and quill.

It was necessary to condition the pad for 5 minutes prior to polishing with phase I slurry. This was done to avoid dishing of copper lines caused by excess material left from a previous copper polish. This same train of thought was used for selecting a 4-min conditioning time for the second slurry. The conditioning time might be in excess of what is truly necessary, but it allowed for a stable process. All pad conditioning was done with deionized water.

Table 4: CMP settings

Polish time (min)	variable
Rinse time (sec)	none
Rotation (Hz, rev/min)	7, 35
Pressure (PSI)	2
Slurry rate (sccm)	25-50 for slurry I 25 for slurry II
Oscillation (cycles/min)	10
Quill Speed (rev/min)	40
Pad conditioning time (min)	5 min for slurry I 4 min for slurry II (conditioning : 7 Hz and zero PSI)

For measuring planarization of copper features two pattern densities were studied. They were the oxide/copper line-space pairs of 45/5 and 5/45 microns, respectively. This corresponds to copper pattern densities of 10 and 90% respectively. Most patterned wafers had oxide trenches between 4600 - 4800 Å deep with a conformal film deposition. The dies sites selected for measuring planarization during CMP are shown in Figure 24. These are the center dies with the best film uniformity. This is because they were parallel with the platen rotation during film deposition.

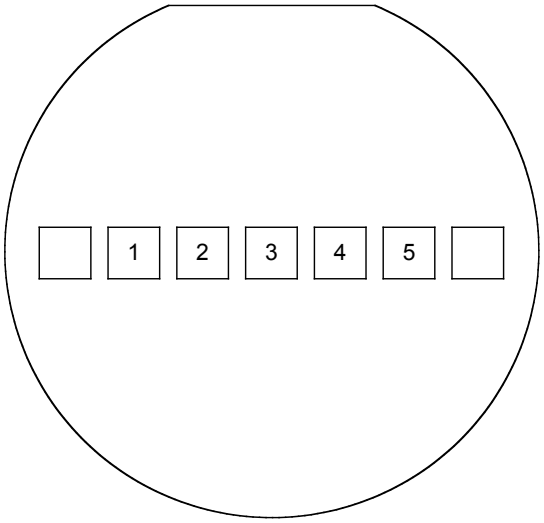


Figure 24: Diagram of selected dies for measuring Cu planarization

Surface topology was quantified by comparing P2 profiles of the features after CMP to those acquired before polish. Further investigations involved using an AFM. The instrument was a Nanospec IIIa from Digital Instruments. AFM was done in tapping (non-contact) mode and the system utilized an optical cantilever.

After CMP, but prior to measurements, wafers were cleaned using a wafer cleaner, model SWC111 from Ultra *t* Equipment Company, that is designed for cleaning after wafer dicing. The tool delivered a high-pressure deionized water spray to remove slurry particles and other particulates. The water pressure was 1200 PSI.

3.4 COPPER MOS CAPACITORS

Cu-MOS capacitors were formed on p-type, silicon (100) wafers by growing a 5000 Å wet oxide and then photolithography patterning them. The oxide was then etched away with buffered HF to form the active regions. The remaining photoresist was then stripped using acetone. Afterwards, a 250 Å gate oxide was grown and 1 μm of copper was deposited with a 100Å TaN/400Å Ta barrier layer. The wafers were polished to remove the copper overburden. The finished capacitors were protected by a resist coating, the backs of the wafers etched bare, and coated with 0.55 μm of sputtered Al to form a back contact. The resist was again stripped using acetone to avoid plasma damage from an O_2 asher. The Al back contact was then sintered in a Heraeus vacuum oven. This was done at 350^{o}C for 60 min at a pressure of 49 mTorr. The vacuum oven was used because copper easily oxidizes at elevated temperatures.

Al-MOS capacitors were also fabricated alongside the Cu ones to act as a control. This was necessary to observe which process steps and conditions might adversely affect the oxide integrity. Full details of the process are available in the Appendix or in chapter 3 of Reference [37]. Aluminum sputtering was performed at 1500 W using an 8-inch target. The sputter and evaporating conditions are listed in Table 5. Only the power density of the TaN deposition for the TaN/Ta/Cu layer is given, because its setting is relevant to any plasma damage that might cause premature capacitor failure. The die arrangement for testing the maximum electric field of the oxide is given in Figure 25.

Table 5: Evaporation and sputtering conditions for MOS capacitors

	Base Pressure (Torr)	**Power Density (Watts/in^2)**
Al evaporation	3.0×10^{-6}	NA
Al Sputtering	5.2×10^{-6}	29.8
TaN Sputtering	1.4×10^{-6}	19.9

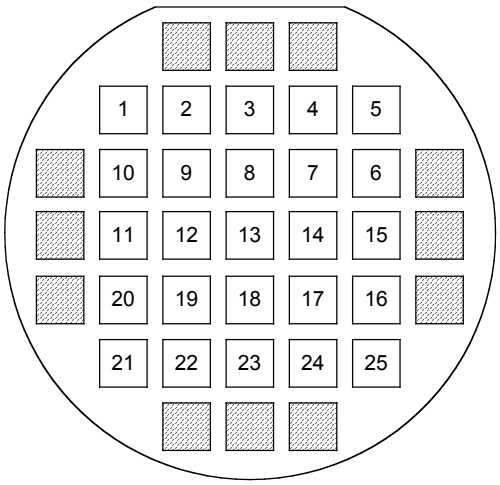


Figure 25: Die arrangement of tested capacitors to determine maximum electric field of oxide

Throughout processing the particle counts on the wafers were monitored using a Tencor 364 Surfscan. The oxide quality was tested using a SemiTest SCA-2000 Surface Charge Analyzer. The oxide thickness was measured using a Nanospec reflectance tool. Finally, capacitor testing was performed using a HP 4145A Semiconductor Parameter Analyzer with a shielded test box.

Testing the capacitors to failure involved applying a negative voltage of -0.5 V to an RC series circuit to place the p-type silicon wafer in accumulation mode. The voltage was then steadily increased by steps of -0.25 V and the current plotted until the capacitors broke down (i.e. the maximum electric field of the oxide was exceeded). Breakdown was signaled by an instantaneous increase in the current. A second sweep was recorded to capture the voltage drop across the resistor. The breakdown voltage was computed from the difference between the voltage at breakdown in sweep #1 and the voltage in sweep #2 (IR loss) at the same current level. This voltage is divided by the oxide thickness to render the electric field value. Moreover, in all tests circular capacitors were used with an area of 8×10^{-3} cm^2.

CHAPTER 4

RESULTS AND DISCUSSION

Stress measurements were performed using silicon wafers with a nominal thickness of 500 μm and a 1 μm of thermal oxide grown on each side. Stress calculations assumed the thickness of the thermal oxide was negligible and the physical properties of the substrate (Young's modulus, Y_s and Poisson's ratio, ν_s) were taken to be that of (110) silicon, $Y_s/(1-\nu_s)$ = 181 GPa.

4.1 TANTALUM FILM STRESS VERSUS THICKNESS

It has been demonstrated through numerous experimental runs that stress in tantalum is a function of film thickness. Figure 26 shows that tantalum films are initially compressive, and as the thickness increases the stress becomes less compressive, and even tensile at some pressures. This phenomenon can be explained by using the grain boundary relaxation model given in section 2.5.

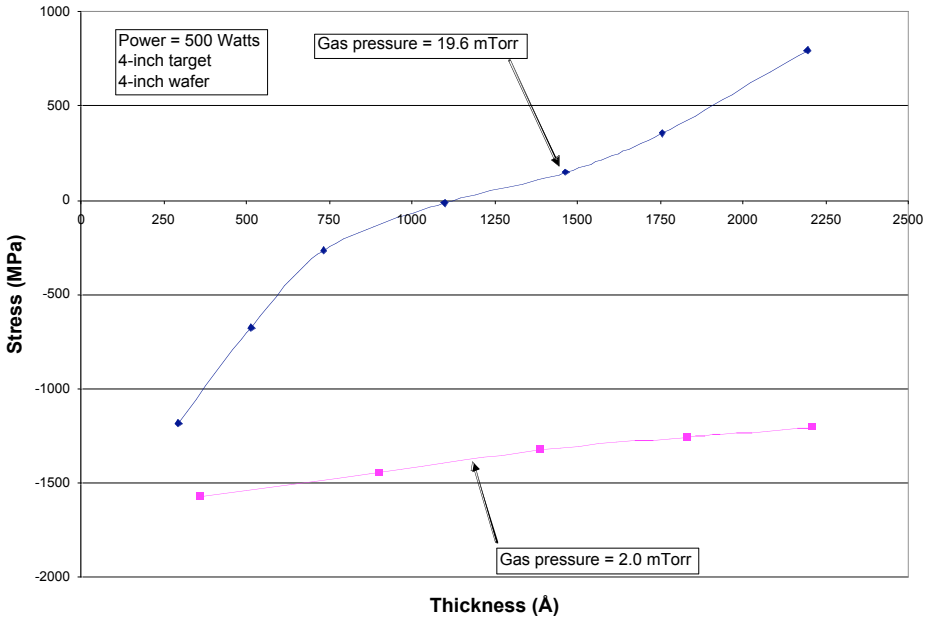


Figure 26: Tantalum stress vs. film thickness for two deposition pressures

Because columnar grain structure results in tensile films a most probable film growth model is given in Figure 27. Initially the crystal grains forming on the substrate are small and as the deposition increases columnar grains are formed. This leads to stronger intergrain forces causing the film to be tensile. One conclusion from this data is that tantalum films used as barrier layers would be in a state of fairly high compressive stress because their thickness is ≤ 500 Å.

Figure 27: Schematic diagram of proposed grain structure change with film thickness [39]

4.2 TANTALUM FILMS STRESS VS. PRESSURE

Figure 28 demonstrates the observed stress-pressure curve for a tantalum film nominally 0.25 um thick. Most notable is the rapid transition region between tensile and compressive stress. In addition, Figure 28 shows that variation in the deposition pressure may result in a wide range of stress values for films of constant thickness.

4.3 EFFECT OF AGING TANTALUM TARGET

Table 6 reveals that tantalum sputtering is complicated by changes in deposition parameters caused by target aging. As the target aged, a decrease in the discharge voltage for a fixed power was observed. Figure 29 shows the stress in these films also changed with voltage. At low/moderate pressures a decrease in voltage caused the film

to become less compressive. At higher pressure a lowering in voltage caused the film stress to be less tensile.

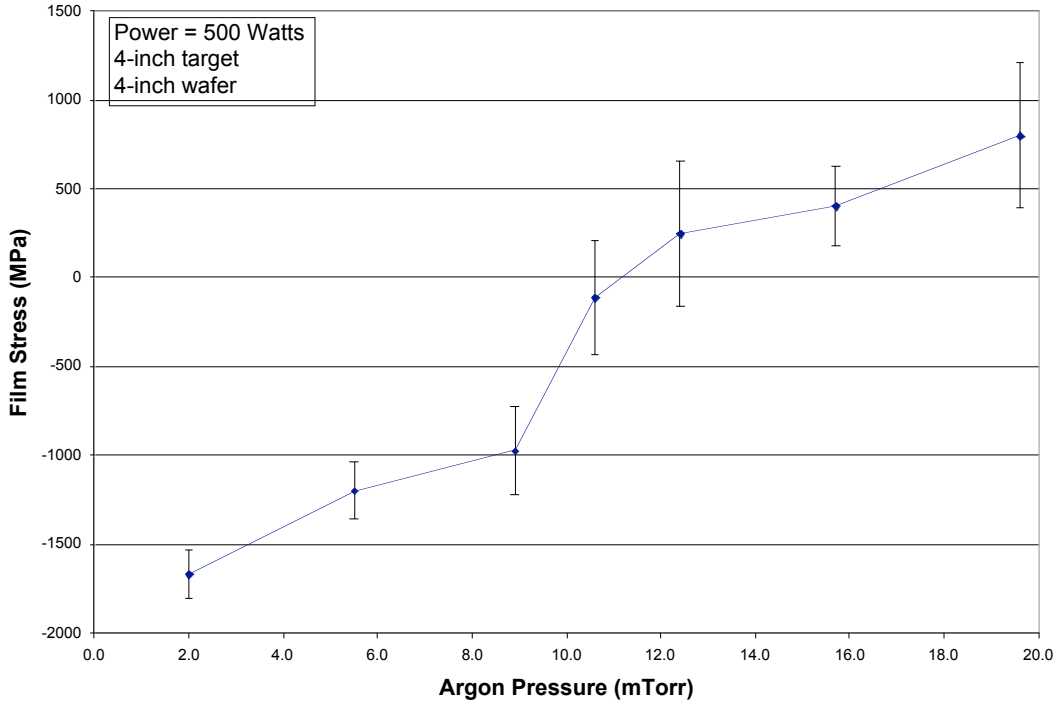


Figure 28: Tantalum stress vs. deposition pressure for a constant thickness. Average thickness is 2408 ± 182 Å

Table 6: Effect of drifting voltage on tantalum film stress. All runs were conducted at 500 W for 15 minutes.

Run	Date	Pressure (mTorr)	Voltage (V)	Stress (MPa)	Dep. Rate (Å/min)
3	8/13/02	2.0	334	-1670	148
26	10/23/02	2.0	296	-1240	165
30	11/4/02	2.0	282	-1120	180
1	8/12/02	8.9	305	-974	163
13	9/13/02	8.9	276	-245	178
9	9/11/02	19.6	285	798	146
25	10/18/02	19.6	264	283	177
29	10/24/02	19.6	258	168	168

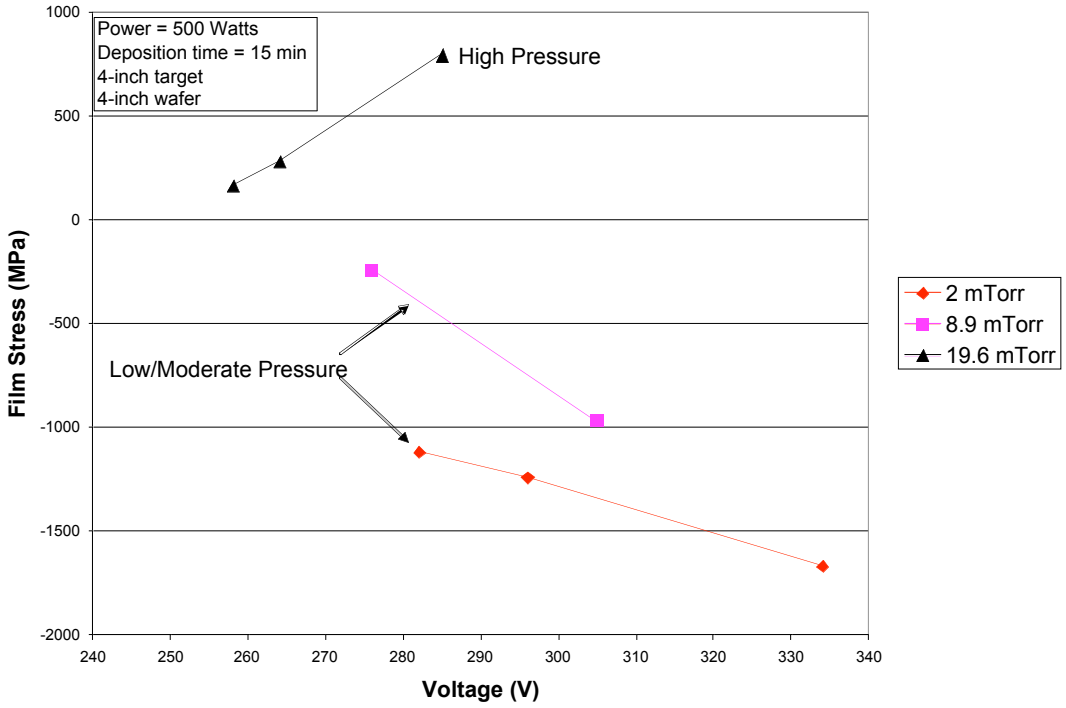


Figure 29: Tantalum film stress as a function of voltage at various pressures

Figure 30 shows the deposition rate as a function of plasma voltage. The higher the voltage, the slower the deposition rate for all pressures investigated. We believe that the increased voltage, or more correctly, the decreased current when power remains constant, is reducing the number of ions striking the target. This in turn reduces the number of neutrals that will impinge on the film due to backscattering or reflection. At low pressures, this will reduce the atomic peening effect that leads to films with more compressive stress. At high pressures, reduced deposition rates will reduce the shadowing effect that leads to enhanced columnar growth and tensile stress. Figure 31 illustrates the role of the reflected neutral particle in sputtering.

Consequently, for greater control of stress it was necessary to sputter the Ta at a constant voltage rather than at a constant power. This results in deposition times that had to be varied to obtain the same thickness for different runs.

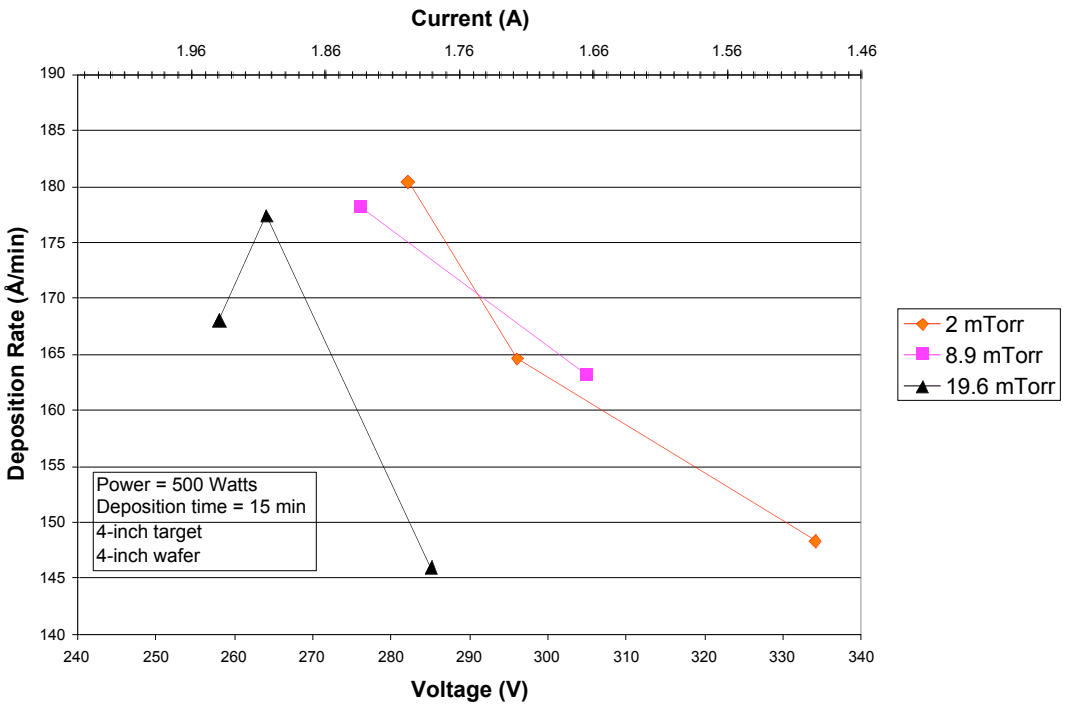


Figure 30: Change in deposition rate with voltage/current at different pressures for tantalum

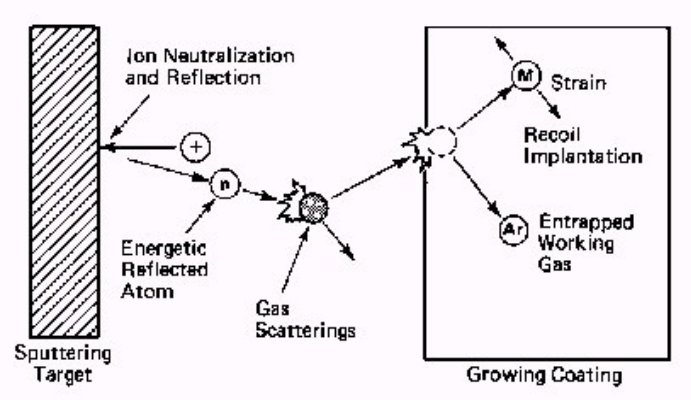


Figure 31: Schematic diagram of a reflected neutral effecting the strain/stress of a sputtered film [24]

Experimental runs using the aged target resulted in a characteristic stress-pressure curve that became linear as seen in Figure 32. When Figures 28 and 32 are superimposed and displayed, as seen in Figure 33, the change in stress becomes clearer. The reduced voltage causes the film at low/moderate pressures to be less compressive. Conversely, the lowered voltage at high pressures causes the film to be less tensile. Altogether, this induces a curve that is more linear in nature.

Therefore, if one operates their process using a fixed power setting, and the voltage is allowed to float, then films deposited earlier in the targets life will not exhibit the same stress levels as films deposited later in the targets life. Consequently, a process that yielded films with good adhesion and a specific CMP removal-rate may eventually result in films that delaminate or have a different removal rate. We attribute the change in plasma voltage to the formation of the racetrack on the target surface and the resulting loss of surface planarity. Since the target is an equipotential surface, a change in its shape will alter the field lines and the angles that the ions strike the target.

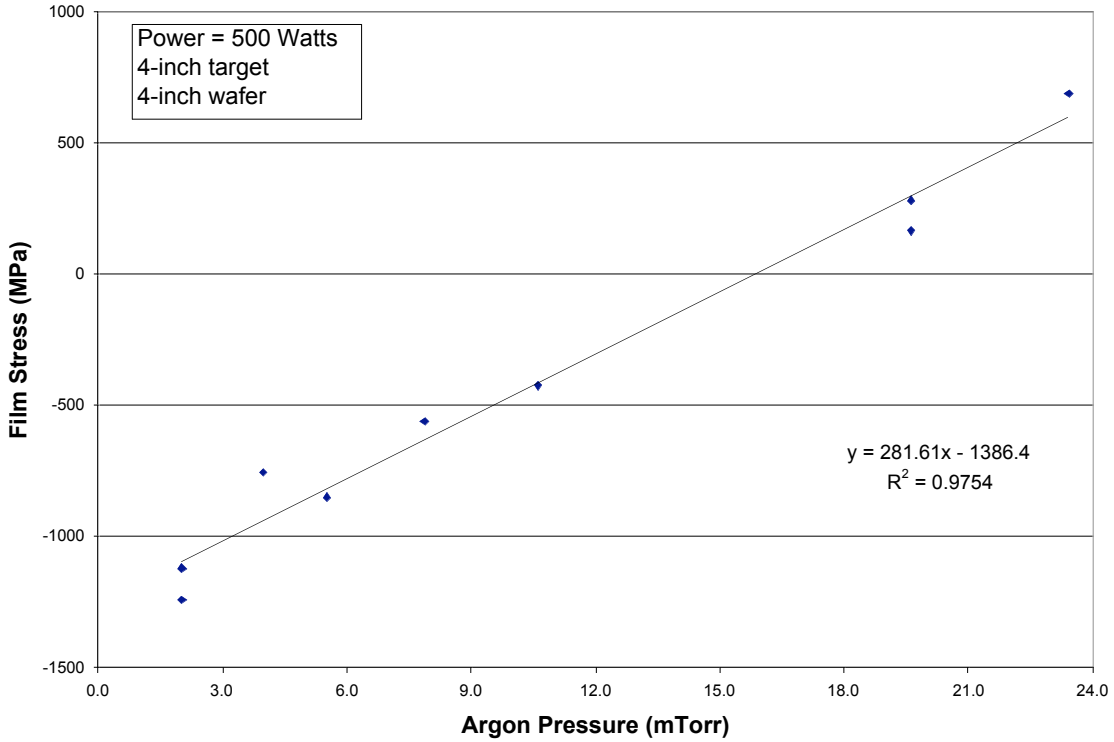


Figure 32: Tantalum film stress vs. pressure from later experimental runs. Average thickness is 2466 ± 186 Å

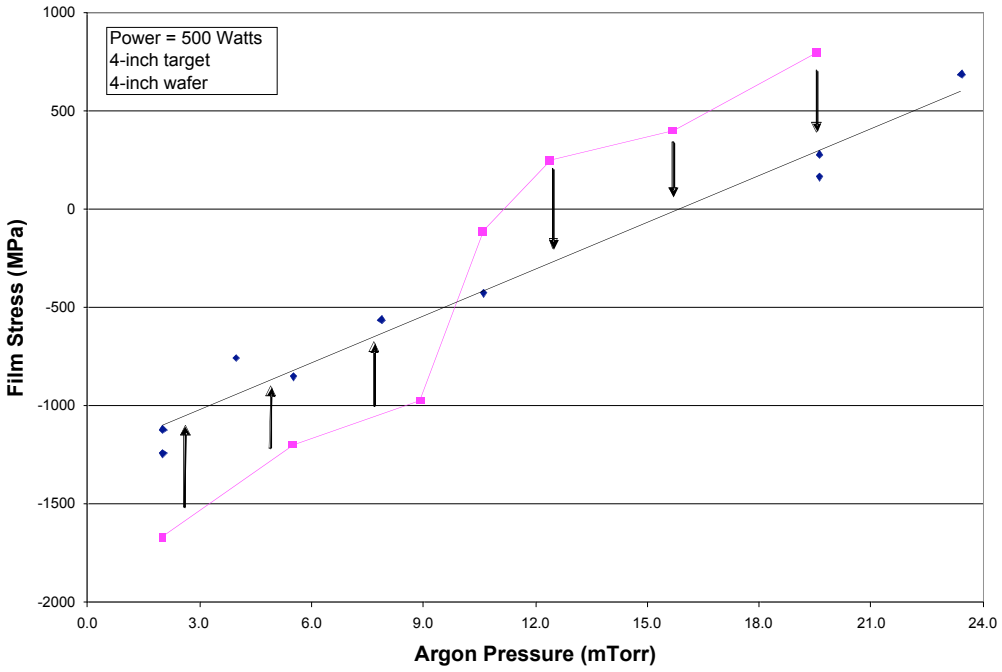


Figure 33: Comparison of tantalum stress-pressure curves (Figures 28 & 32)

4.4 TANTALUM STRESS GRADIENT

Whenever tantalum is deposited a large standard deviation (STDEV) in stress was observed. In fact the percent covariance, defined as [100% (STDEV) / Average], averaged over thirty-one runs was 86%. This may be caused by a stress gradient in the film that is due to film thickness non-uniformity across the wafer. This may be due largely to the fact that a 4-inch target was being used to sputter tantalum onto 4-inch wafers. Ideally, the target should be an 8-inch diameter for better uniformity. A general trend, when going from the center toward the edge on a wafer, was to see the thickness decrease. In fact, in some cases we obtained both tensile and compressive stresses on the same wafer. However, when the average stress was plotted the observed trends were very reasonable and consistent with literature findings (compare Figures 14 and 28).

4.5 STRESS STDEV VS. STRESS FOR TANTALUM

Using thirty experimental runs, a curve of stress standard deviation vs. stress was constructed and is displayed in Figure 34. It shows a general trend where films having higher levels of compressive stresses have a smaller standard deviation. Even though the R^2 value is only 0.67, it is fairly significant because the curve encompasses all experimental data, which do not necessarily share common deposition conditions. In addition, variation in film uniformity or wafer thickness was not taken into account for all samples and that would tend to randomize the data. However, one may conclude that for reasonable uniformity in stress values for these Ta films, deposition parameters should be set to obtain films having compressive stress.

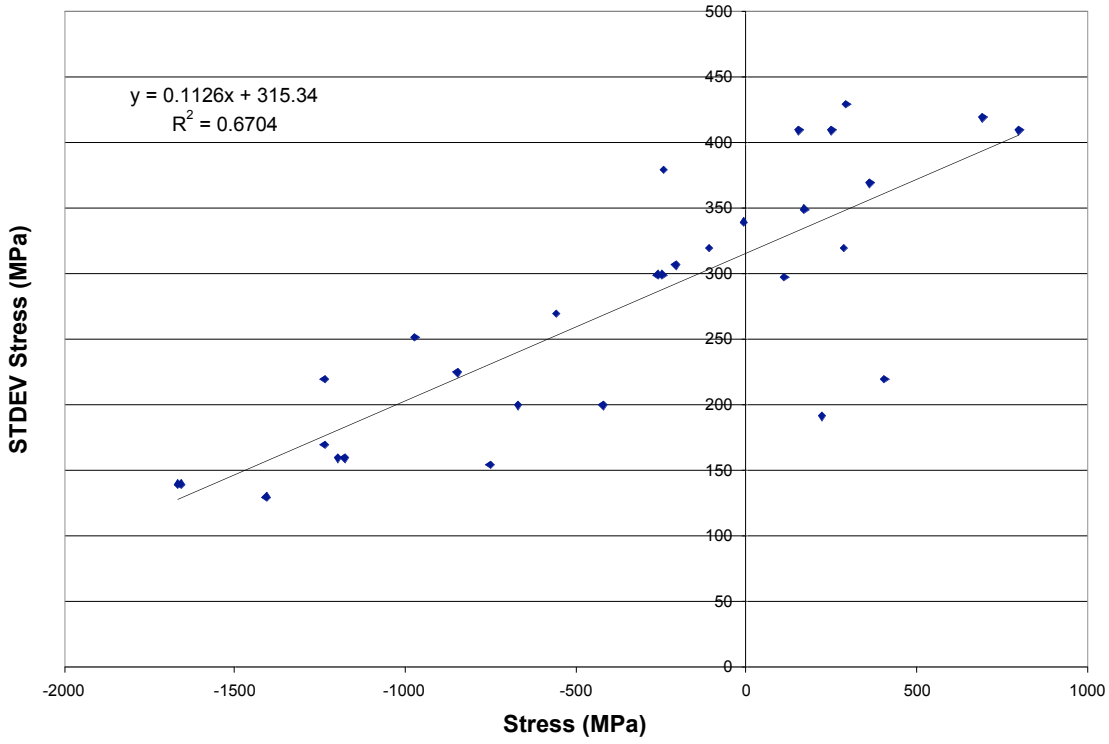


Figure 34: STDEV stress vs. stress for tantalum

4.6 TANTALUM CMP

Figure 35 shows the results of a CMP study of removal rate versus stress for select Ta films. A substantial difference in CMP removal rates of tantalum occurred between approximately -400 and -1200 MPa of compressive stress when polishing with EKC phase 1 copper slurry. Recall, phase 1 slurry should stop on the Ta layer. It was found that for high levels of compressive stress in the Ta, the CMP process removed the Ta and the film did not act as a polish stop. If the removal rate is too low, there may be dishing of the copper. Therefore, the ability to control the stress in the Ta film appeared to be critical to the CMP process as removal rates of the barrier layer can effect the overall performance.

Quantifying removal rates for films having large values of compressive stress proved difficult. After polishing a film with high compressive stress, a lower resistivity was found with 4-point probe measurements. This may be due to a change in the material's crystal structure or a highly conductive tantalum complex that formed after polishing. Consequently, exact removal rates could not be determined.

Tseng et al. showed that for aluminum and tungsten the CMP process increases the resistivity of the metal after polish by randomizing the crystallographic orientation [25, 26]. In addition, they demonstrated the correlation between microstructure of a metal thin film and its CMP performance. Metal films with a smaller grain size polish faster. This can be attributed to the larger grain boundary area available for slurry attack, accelerating the chemical erosion process and hence CMP removal rate.

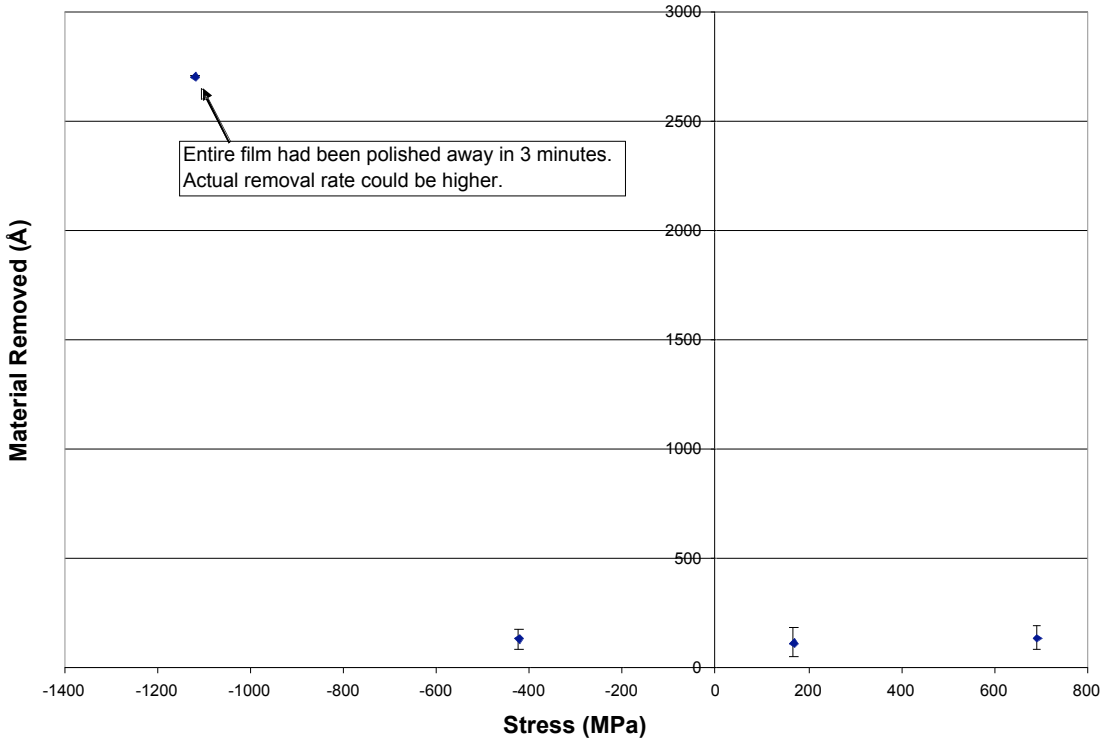


Figure 35: Tantalum CMP removal vs. stress for a 3 minute polish. Phase I slurry, press = 6 Psi, pad rotation = 9 Hz (45 rev/min), quill speed = 55 rev/min.

4.7 CHEMICAL ETCHING OF TANTALUM

A short study was carried out to investigate the chemical portion of CMP. Using an aqueous solution of 29% HF and 29% HNO_3 , Ta films of various stress levels were etched. The trend for the chemical removal rate of tantalum was that a more tensile film etched faster. This is opposite to what was found for the CMP removal rates. This suggests that the CMP process is more mechanical in nature. This is also supported by the fact that there was a zero static etch rate for tantalum in EKC phase II slurry that is used to remove the barrier layer. Figure 36 summarizes the chemical etch rates observed.

Both these observations may be the result of tantalum forming a strong passivation layer. Tantalum is known to readily form a hard, protective oxide (Ta_2O_5) in aqueous solutions that makes it very inert [27, 28]. This film may have to be

mechanically removed, and the chemistry then creates a fresh passivation layer. At this point we concluded that the high removal rate of Ta films having high values of compressive stress is due to enhanced mechanical mechanisms.

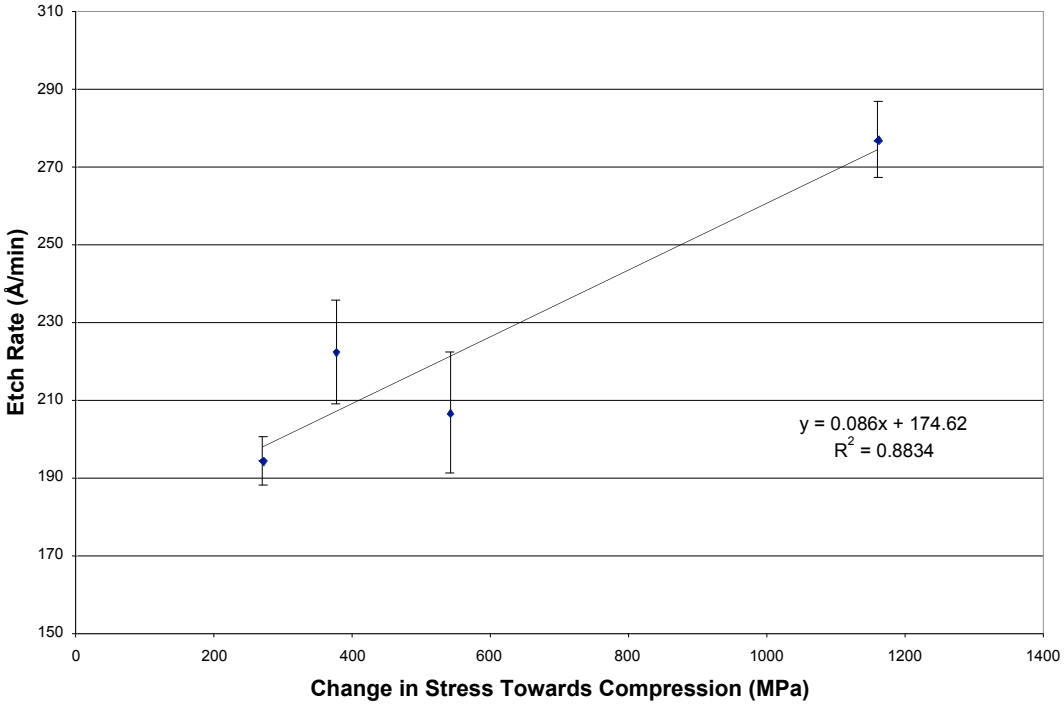


Figure 36: Chemical etch rate vs. change in stress for tantalum. A solution of 29% HF and 29% HNO_3 was used to etch tantalum for 3 minutes. When the film is etched it becomes more compressive.

4.8 COMPARISON OF TANTALUM ETCHING AND CMP RESULTS

An interesting result in Figure 36 is that as a tantalum film is chemically etched, the remaining film became more compressive. Therefore, it is possible that when a film is being polished the removal rate could increase as the film becomes increasingly compressive with decreasing thickness. This assumes that the film stress is compressive enough to give and increased removal rate (see Figure 35). It is also possible that the polish rate could increase when a critical stress is met. If this is true, the downward force

of the polishing head would become a critical parameter in the control of the polish and may require dynamic adjustments during the process as the film's stress changes.

4.9 COPPER/TANTALUM ADHESION FAILURE

Table 7 shows adhesion results that were observed for 1.4 μm thick films of sputtered copper on tantalum films having either tensile or relatively high values of compressive stress. Only the run having a Ta film with -1410 MPa of compressive stress had the Cu layer adhere (passed scribed tape test). It appears that some value of compressive stress less than ≈ -1500 MPa is required to achieve good adhesion with tantalum alone. This data underscored the need for tantalum films with the compressive stress controlled. The next step was to determine the effects of a TaN barrier layer to eliminate possible delamination at the barrier layer/oxide interface.

Table 7: Cu-Ta bilayer scribed tape test results

Run	Ta Film Stress (MPa)	Cu Film Stress (MPa)	Passed Scribed Tape Test
1	-1660	20.0	No
2	-1410	20.0	Yes
3	Tensile[1]	Tensile[*]	No

4.10 TANTALUM NITRIDE

The first step in developing a tantalum nitride process for improved adhesion was to characterize its hysteresis behavior, if any. This was necessary to determine the working range of N_2 flows for deposition of the film. The plot of sputter voltage versus N_2 flow rate is given in Figure 37. In this plot there are two sets of data. The earlier data showed a trend of increasing voltage for increasing N_2. This indicates that the target is not poisoned, and that the nitrogen is either being incorporated into the film or onto the target surface. Repeating the study at a later date resulted in the second set of data. Here the voltage saturated at N_2 flow rates above 25sccm. This is indicative of a saturated

[1] Data are from older results, exact stress values are not known.

target surface. It was not readily apparent as to the cause(s) of the observed changes. We did not pursue this topic.

Figure 38 shows that chamber pressure decreased with increasing N_2 flow. This indicated that the deposited film was capable of incorporating an ever-increasing amount of nitrogen. At a nitrogen flow rate of 30% (15 sccm), the tantalum nitride films became dielectric in nature, and their resistivity was unable to be extracted from 4-point probe measurements.

Data for resistivity vs. percent nitrogen is shown in Figure 39. It shows the resistivity was low and fairly constant until approximately 15% N_2. Therefore, it was decided to run the tantalum nitride process within this range at 7.5 sccm (15% N_2 by flow rate). Figure 40 shows that the deposition rates for TaN were fairly constant between 84 and 92 Å/min for $\leq$ 12 % N_2 (flow rates of $\leq$ 6 sccm). These fluctuations were most likely due to relatively poor control of flow because we were at the low end of the mass flow controller's range. At the 12% setting, Figure 41 shows that the tantalum nitride resistivity decreased with increasing film thickness. The high values of resistivity observed in the initial stages of film growth are most likely due to scattering of the conducting electrons at the film surface (Fuchs-Sondheimer effect) during 4-point probing [33]. Thinner films tending to have smaller grains that can lead to increased grain boundary scattering of electrons. In addition, it may be necessary to have several hundred angstroms of material to make the film continuous.

In Figure 40 the trend of deposition rate verses percent nitrogen is opposite to what is found in the literature [29, 30, 31, 32]. The deposition rate is known to actually decrease with increasing N_2. However, others have indicated that under the same conditions the growth mechanism of tantalum nitride films depends greatly on the deposition system such as power density, total ambient gas pressure and target-to-substrate distance [30]. Consequently, different research groups fail to publish consistent results.

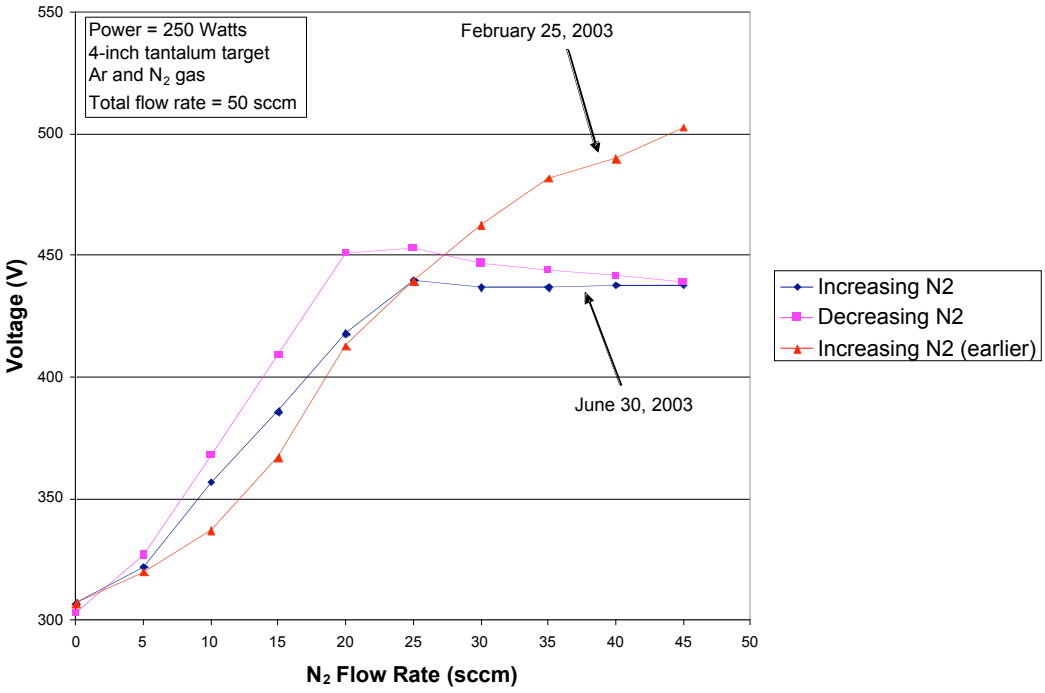


Figure 37: Hysteresis behavior of tantalum nitride deposition

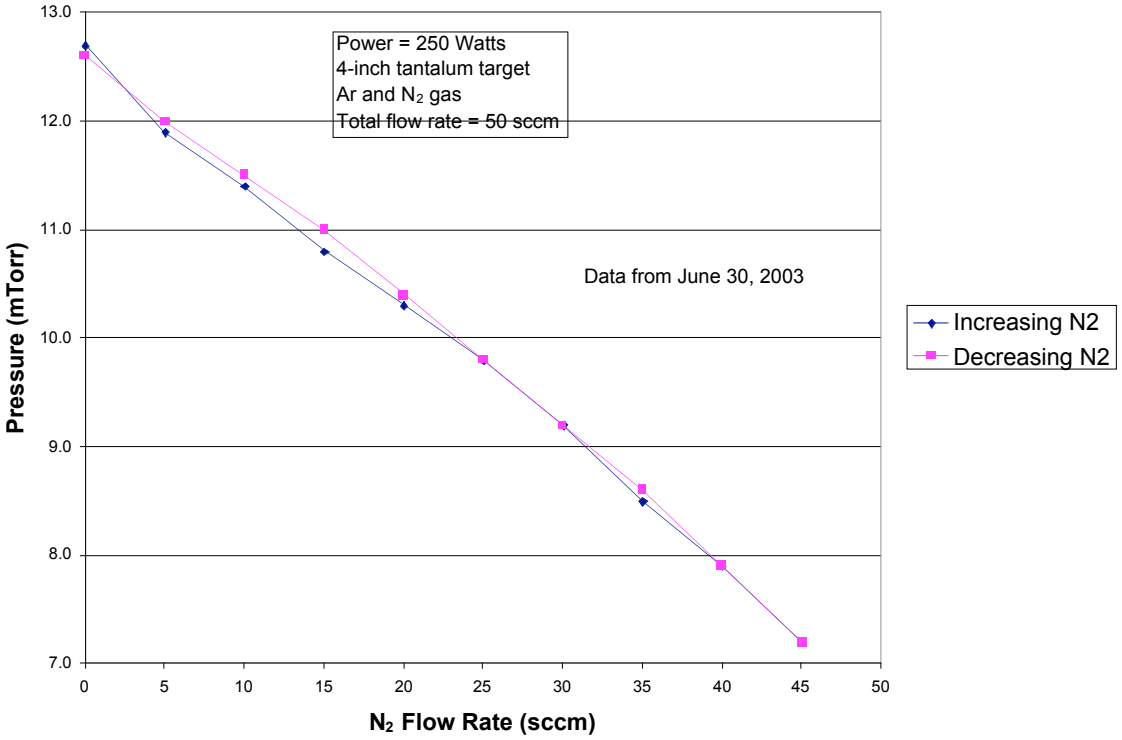


Figure 38: Deposition pressure vs. N_2 flow rate during tantalum nitride hysteresis

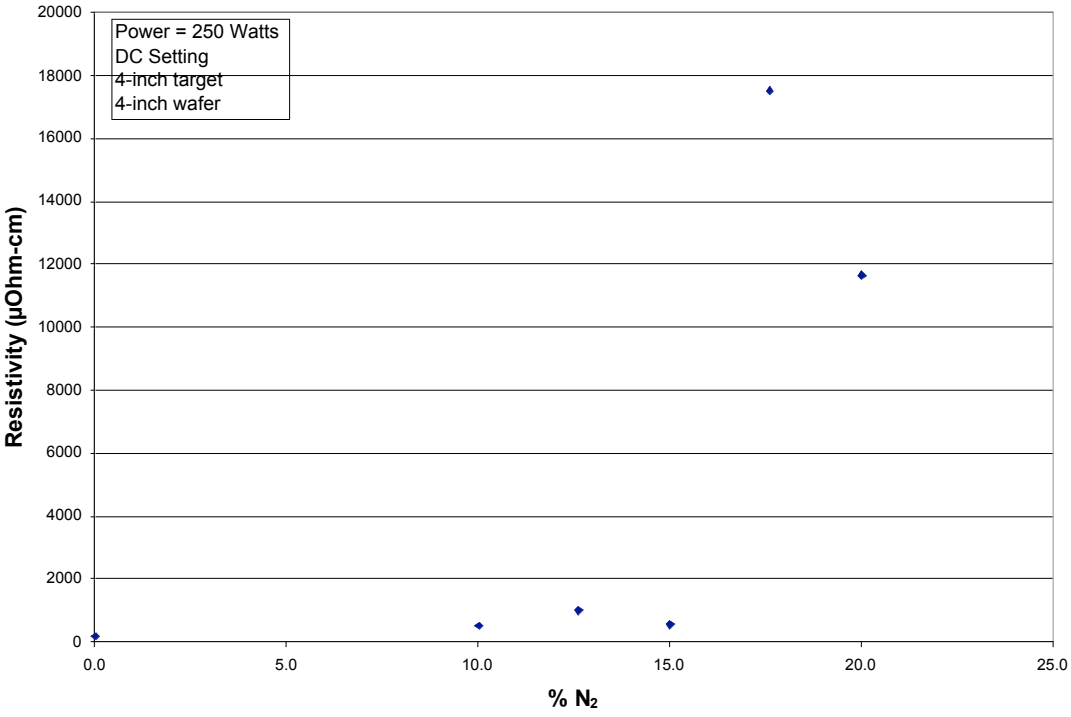


Figure 39: Tantalum nitride resistivity vs. percent N_2

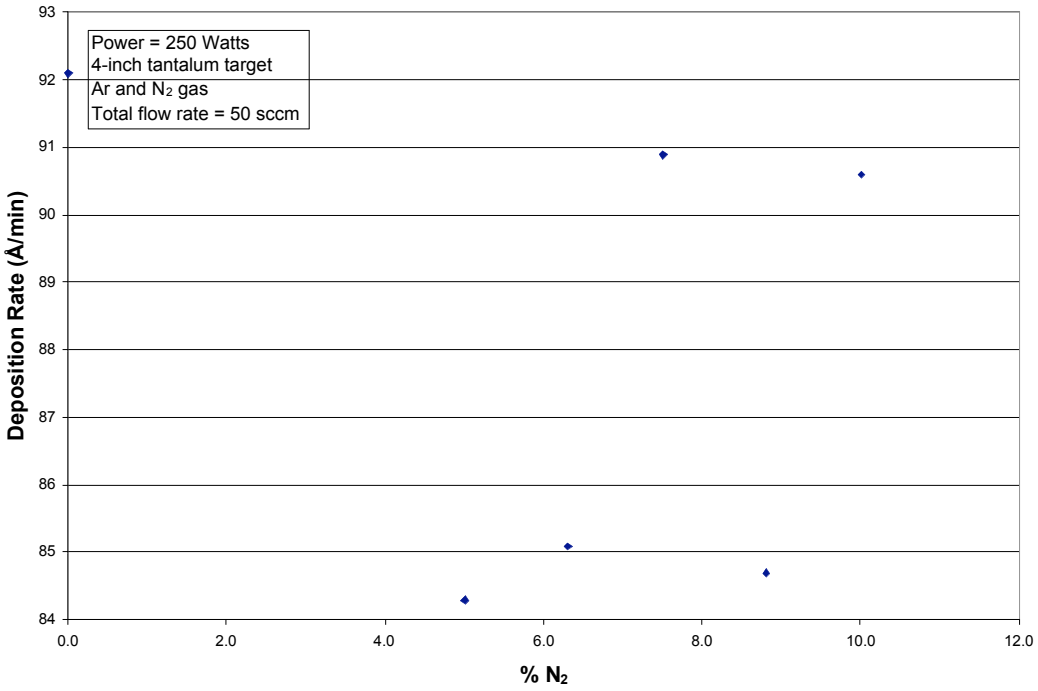


Figure 40: Deposition rate for tantalum nitride vs. percent N_2

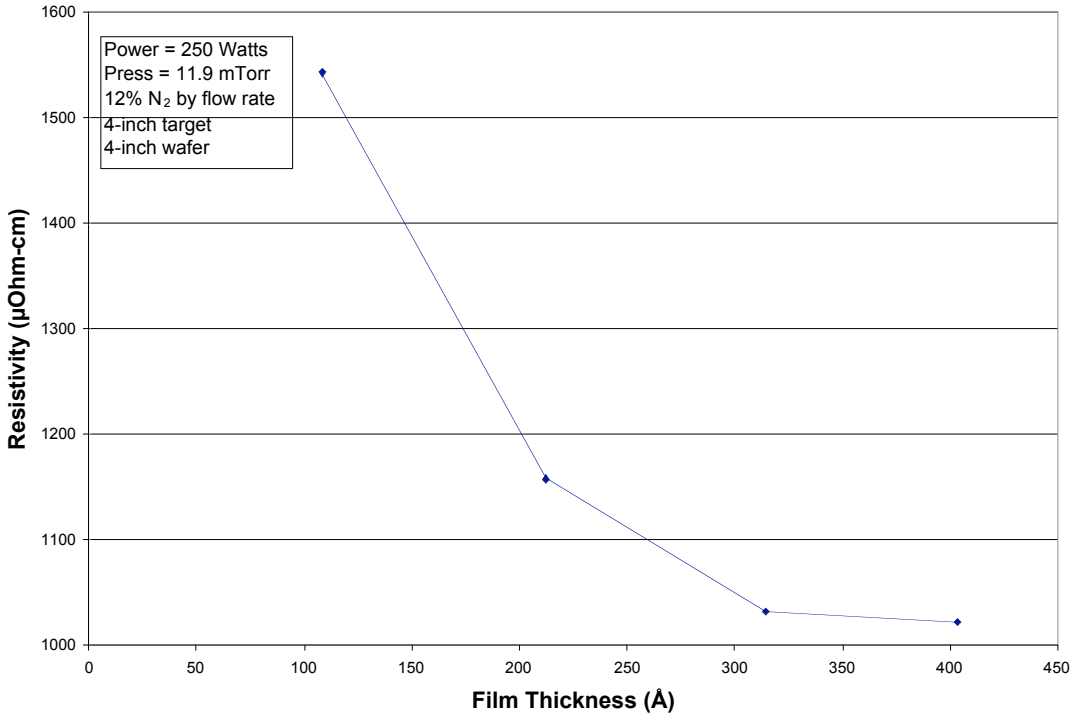


Figure 41: Resistivity vs. thickness for tantalum nitride films

4.11 TANTALUM DEPOSITION

Because of the observed aging effects and the need to control the stress in the Ta film, the decision was made to sputter Ta for the TaN/Ta film stack at a constant voltage. Depositions were performed at 300 V. Figure 42 shows how the deposition rate changed with discharge current. This curve was used to engineer the thickness of the tantalum films by adjusting the time once the current was known.

During the sputter run the discharge current would vary to some extent. To reduce the variation a tantalum presputter was done for 10 minutes where immediately after, the film was deposited.[2] Figure 43 demonstrates that film uniformity was better at lower pressures. This would be expected since target atoms are arriving at the substrate with fewer random collisions.

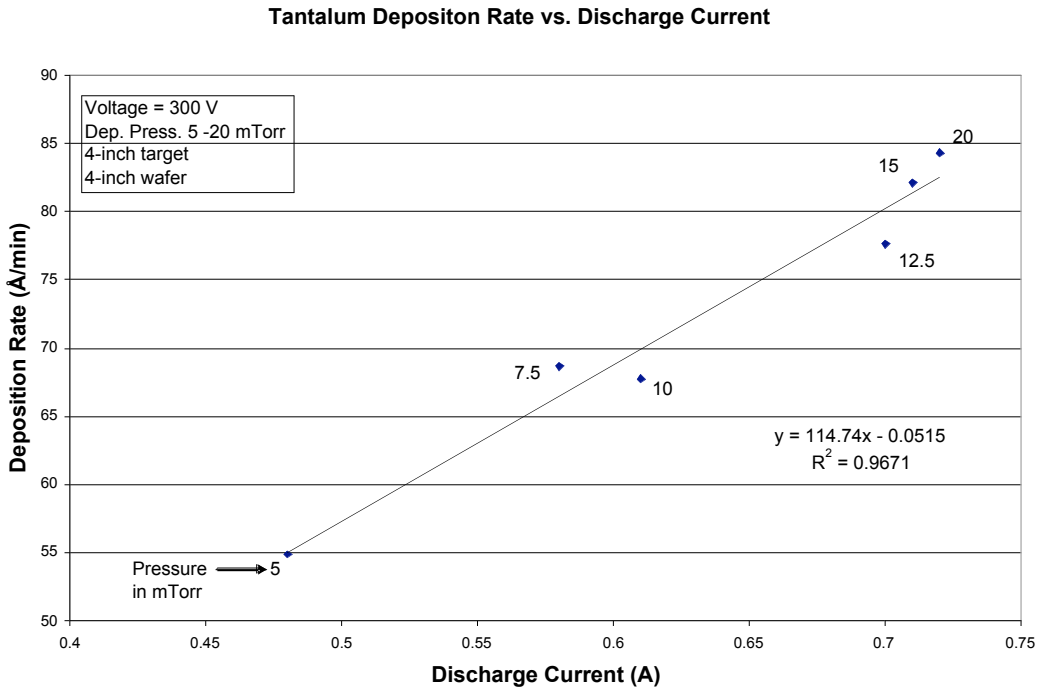


Figure 42: Deposition rate vs. discharge current for Ta sputtering performed at constant voltage. Deposition pressure is indicated next to data points on graph.

[2] A 5 minute presputter was not sufficient to help stabilize the discharge current. Therefore, a 10 minute conditioning was used.

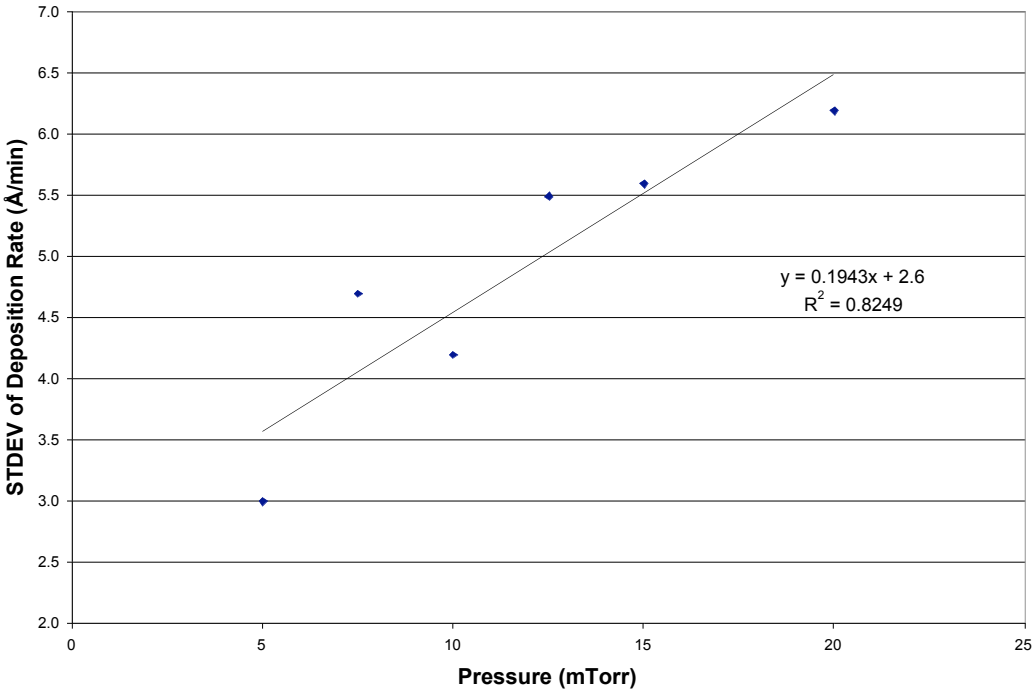


Figure 43: STDEV of tantalum film deposition rate vs. pressure. Measurements were taken at positions X3-X7 (40-mm center portion).

4.12 TaN/Ta FILM STACK

Reactive sputtering was used to produce a 500 Å thick TaN/Ta film stack. As illustrate in Figure 44, the stress varied with thickness for the TaN layer. Consistent with previous results, an increase in TaN thickness resulted in films with less compressive stress. However, the TaN films were observed to have a much lower level of compressive stress, as compared to a similar thickness of Ta. Deposition of Ta on top of TaN, with the thickness of both layers adjusted to produce a total film stack thickness of 500 Å, resulted in the stress of the film stack being fairly constant at approximately -690 MPa with little variation for different combinations of TaN/Ta thicknesses. This demonstrates that the film stress exhibits little variation using the bilayer approach. In fact, every combination of TaN/Ta passed the scribed tape test after copper was deposited on it.

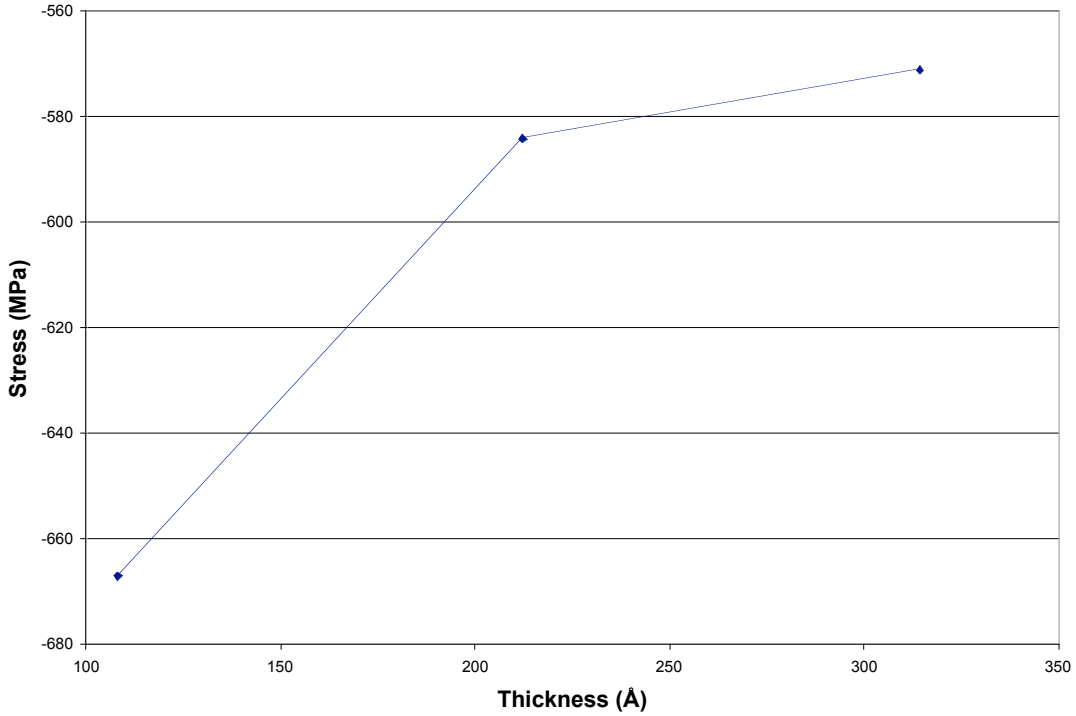


Figure 44: Stress vs. TaN thickness deposited on oxide

4.13 X-RAY DIFFRACTION ANALYSIS

Three samples were analyzed by XRD. Table 8 shows the sample IDs. The full XRD report is in the Appendix, but the key results are summarized here. The diffraction spectrum for Sample A1, shown in Figure 45, revealed that the reactively sputtered TaN had a 1:1 stoichiometry and was amorphous. This is advantageous because the interface fracture energy increases with nitrogen content up to the 1:1 stoichiometry [11]. Moreover, it exhibits superior barrier layer properties over other TaN complexes (see Table 2). Sample A2 produced the spectrum in Figure 46 that shows the presence of highly orientated β-Ta, with the (00L) planar orientation. This was expected since only the tetragonal form of Ta is known to form on silicon dioxide when sputtering. Sample A3 shown in Figure 47 reveals that highly orientated β-Ta is formed as well when sputtered on TaN. By using the Scherrer technique the nominal crystal size for Ta in samples A2 and A3 was determined to be 230 and 160 Å, respectively. The thicker sample had larger crystal grains. This is expected because grain size tends to increase

with film thickness. These results confirmed that the deposition parameters selected for the TaN/Ta film stack produce the film properties desired for CMP applications.

Table 8: Samples analyzed by XRD

Sample ID	Film on SiO_2/Si
A1	400 Å TaN
A2	630 Å Ta
A3	100Å TaN/400Å Ta

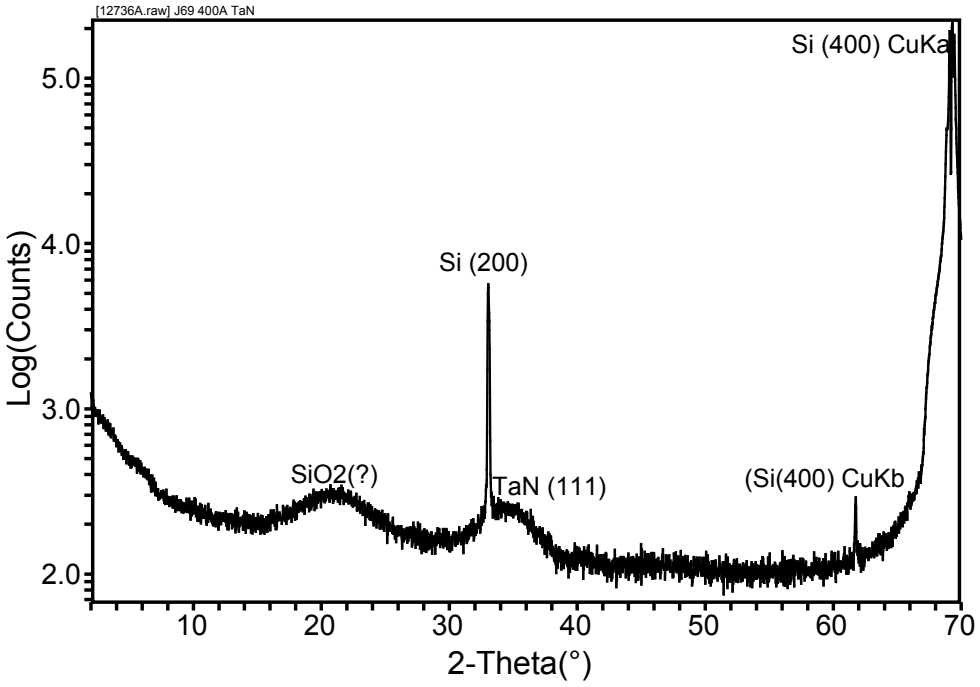


Figure 45: XRD spectrum of 400 Å TaN film, sample A1

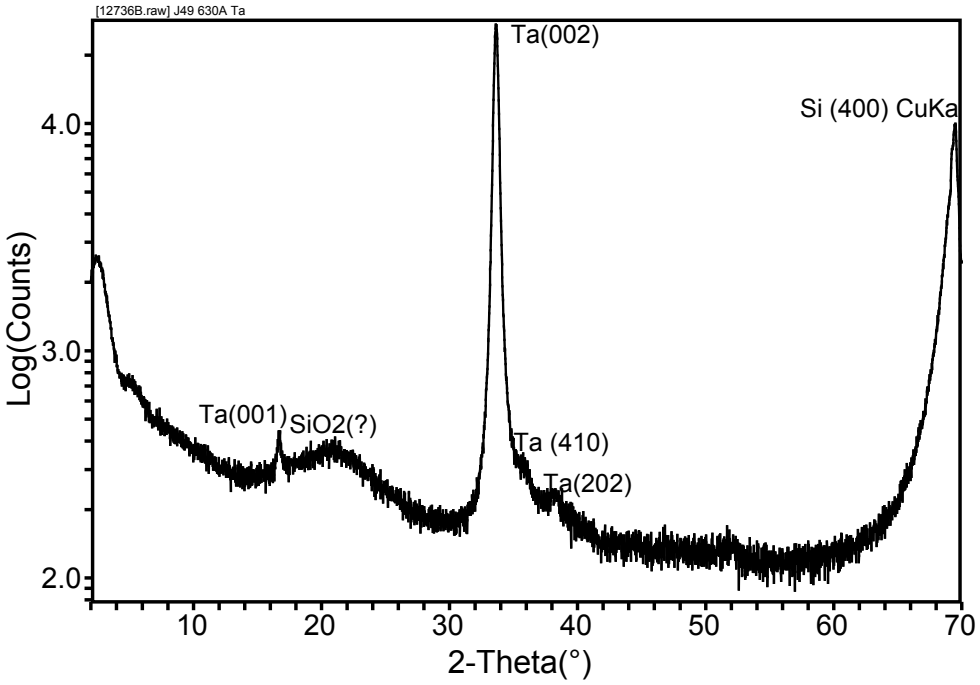


Figure 46: XRD spectrum of 630 Å Ta film, sample A2

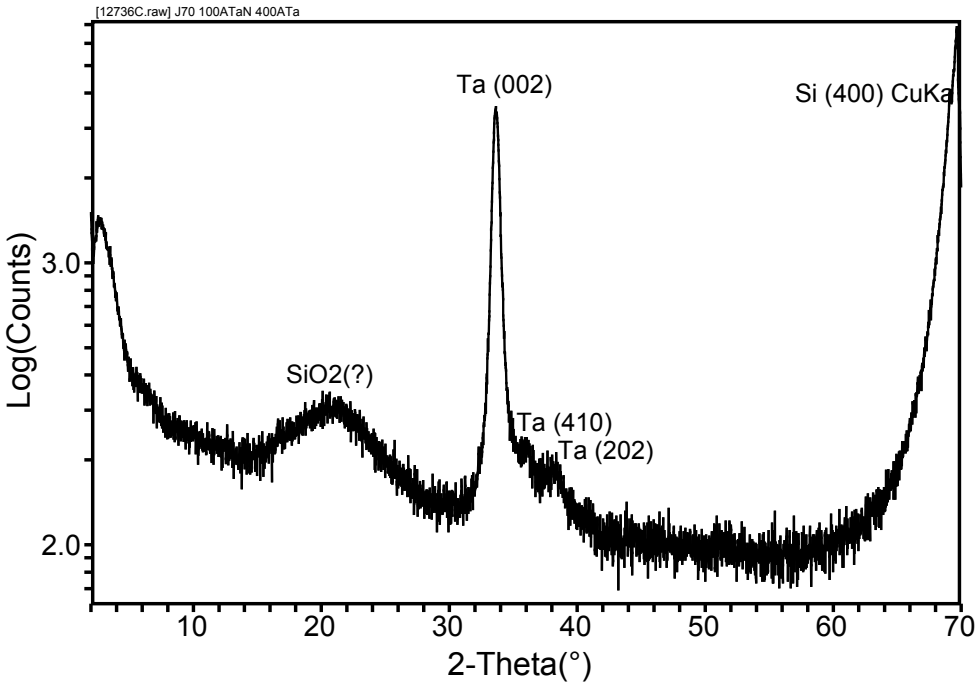


Figure 47: XRD spectrum of 100 Å TaN/400Å Ta film stack, sample A3

4.14 COPPER DEPOSITION

A barrier layer of 100Å TaN / 400Å Ta was chosen for use with the deposition of copper. With this combination the Ta film had the lowest stress and theoretically would have the best adhesion. Copper was sputtered at 1500 W to produce a 1 μm film with a resistivity of 2.02 μΩ_cm as measured with a ResMap tool. This compared favorably with the bulk value of 1.7 μΩ_cm resistivity for copper. In addition, the copper stress was 27 MPa of tensile stress, which was observed to relax by 1-2 MPa after a week's time. The change in stress can be contrasted with that of Ta, which showed no change in stress after 5 weeks. In Table 9 a copper deposition run sheet is given. The settings are nominal and the use of constant voltage for the Ta sputter is to correct for target aging. Changes in the Cu deposition over target lifetime were not undertaken in this work as the 8-inch target ages much slower than the 4-inch target, and the TaN/Ta adhesion was the more important issue.

Table 9 : TaN/Ta/Cu log sheet. In the far-right column of the table the values for volts, power, current and argon flow are not listed because they are floating values and therefore are run dependant.

Time (sec)	20	HF dip (100:1, H_2O:HF)
Pressure (Torr)		40 min bake at 250°C after going into high vacuum. Wait at least 20 min after bake step.
Tantalum Presputter	4" Target	
Power (W)	250	
Time (min)	5	
TaN Presputter	4" Target	
Power (W)	250	
Time (min)	5	
TaN Sputter	4" Target	I = amps, V = volts
Power (W)	250	
Time (sec)	70	
N_2 Partial Pressure	12 %	N_2: 6 sccm, Ar: 44 sccm
Pressure during sputtering (mTorr)	11.8 – 12 mTorr	
Tantalum Presputter	4" Target	
Voltage (V)	300	
Time (min)	10	
Tantalum Sputter	4" Target	I = amps, P = watts

Voltage (V)	300	
Time (sec)		
Pressure during sputtering (mTorr)	10.0	Ar: sccm
Copper Presputter	8" Target	
Power (W)	1500	
Time (min)	2	
Copper Sputter	8" Target	I = amps, V = volts
Power (W)	1500	
Time (min)	20.33	1220 sec
Pressure during sputtering (mTorr)	5.5	Ar: sccm

4.15 COPPER CMP

Blanket copper wafers were polished using the two phases of the EKC slurry. The different phases are used in a two step process. See Section 3.2 for more details. The removal rates of copper and other films are listed in Table 10.

Figures 48 and 49 are profiles of polished oxide/copper line-space pairs of 45/5 and 5/45 microns, respectively, which constitute Pattern Factors (PF) of 0.1 and 0.9 for percent Cu. The 5/45 features show substantial rounding of the oxide from polishing. This is because there is less SiO_2 area to support the force, and the increased localized pressure causes an increase in the SiO_2 polish rate [34]. The thinning occurred within the first few minutes of CMP when there was still an overburden of copper. As a result of this once the edges touch, the feature is rapidly polished away and this pattern density planarized faster than the 0.1 PF features. Figure 50 illustrates this feature after polish.

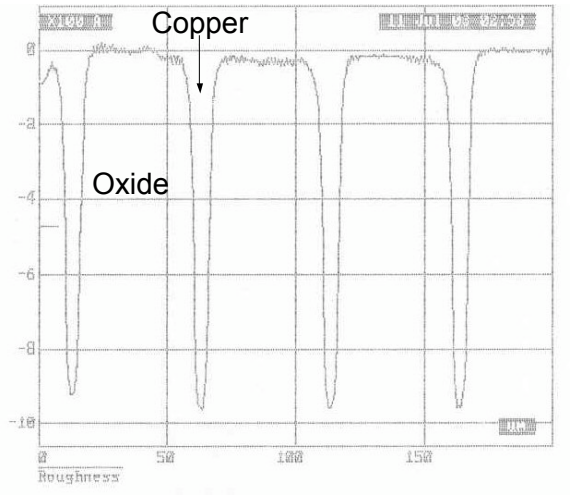


Figure 48: Profilometer scan of oxide/copper features after phase I slurry polish (45 µm oxide/5 µm Cu feature). Sample still had barrier layer.

Table 10: Blanket film removal rates at 2 Psi and 7 Hz (35 rev/min)

Film	EKC phase I slurry (Å/min)	EKC phase II slurry (Å/min)
Cu	900	33
Ta	22	440
TaN	NA	660
Oxide	NA	89

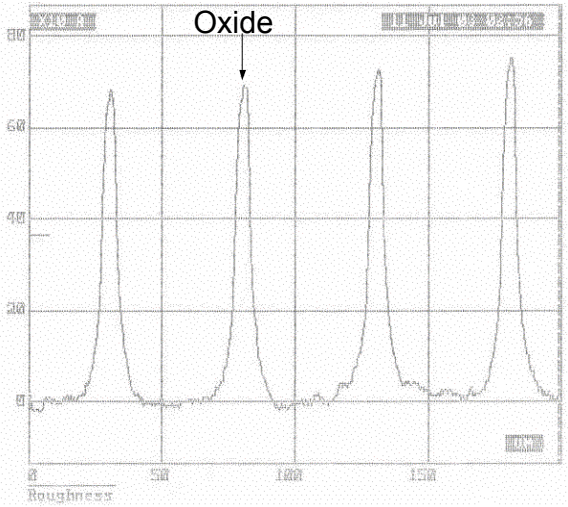


Figure 49: Profilometer scan of oxide/copper features after phase I slurry polish demonstrating oxide thinning (5 µm oxide/45 µm Cu feature). Sample still has the barrier layer.

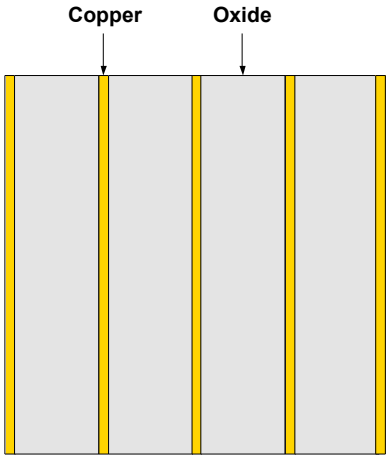


Figure 50: Diagram showing the 45 µm oxide/5 µm Cu features that were used to characterize the CMP process

When exposing wafers for CMP the die was replicated in a 7x7 array. However, one the wafers was patterned with only a 5x5 array. This slight change in the number of die on the wafer made a significant alternation in overall wafer polish rate. Figure 51 demonstrates this effect. For the smaller array copper was left in the unpatterned areas even though both wafers were polished for the same amount of time. The lack of topography decreased the localized pressure and resulted in lower polishing rates.

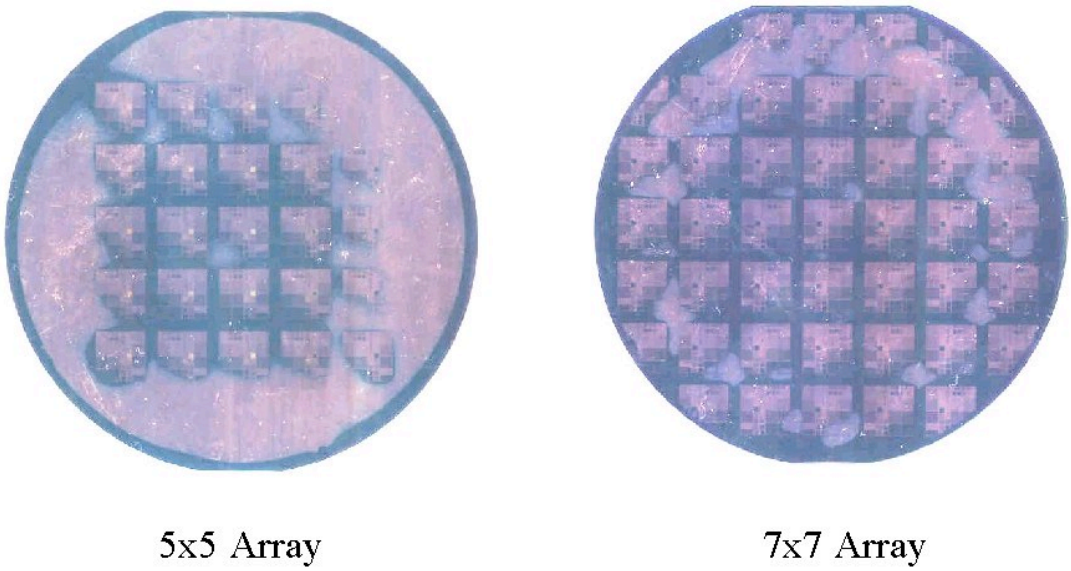


Figure 51: Different wafer polishing characteristics for wafers under the same conditions but with a different number of die. The final polish for each wafer was performed with the second slurry. Note that any copper left after phase I slurry will most likely remain after the second phase slurry because it has a very low copper removal rate.

Figure 52 shows the characteristic planarization curve developed for the copper damascene process. In the Figure the Remaining Step Height (RSH) is the step height after polish normalized to the original step height. The curve was constructed by selecting the desired pattern density and polishing several wafers for different lengths of time. The same wafer was not repolished after the step height was measured because successive polishes will not yield the same result as a continuous polish for the same amount of time. One assumption for this curve is that your wafers have good uniformity in starting topography and film thickness.

The result is that the copper planarization proceeds linearly until there is no remaining step height. Afterwards, copper features begin to dish and the step height will increase. At this point, you cannot recover planarity unless one were to develop a slurry for oxide removal with a Cu polish stop.

This result is key to a successful CMP process. For our 1 um thick film, we observed that it took about 8 minutes to reach the barrier layer for a removal rate of 1250 Å/min. It also shows that it took 6 minutes, or the removal of 7500 Å of Cu, to reach planarity for a 0.47 um feature step. Therefore, if one did not have 7500 Å or more of Cu, one could never achieve planarity. This amount of copper in excess of the given feature step is termed the overburden and is critical in achieving planarity.

.

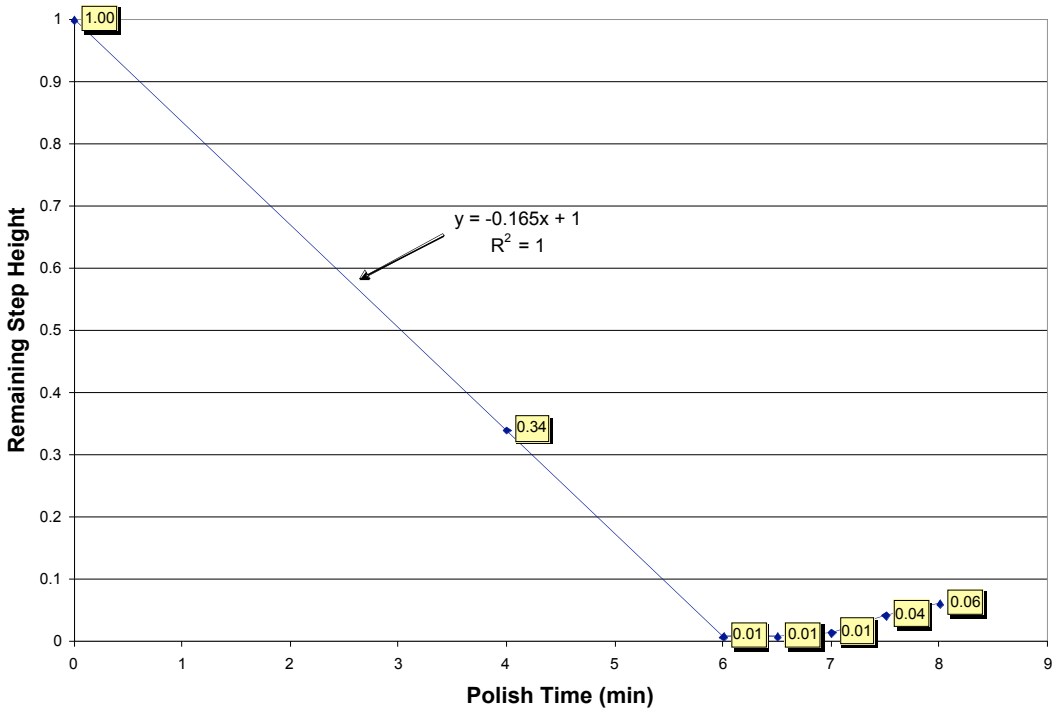


Figure 52: Remaining step height vs. polish time in copper damascene process for a series of wafers being polished for different times

After polishing copper features with the first slurry it is necessary to employ the second one to remove the barrier layer. When doing so, any copper that remains on the wafer is unlikely to be removed because the phase II slurry has a low copper removal rate. Consequently, a slight copper overpolish is necessary to insure most of the copper is removed with the phase I slurry. This is due to uniformity issues of the copper film.

For developing a standard process, a 7.5 min phase I polish was used. This gave approximately three times the number of usable die (27 vs. 9) than when only a 7 min polish was employed. A longer 8 min phase I polish was not used because the step height reduction increased by 2.3 times. Figure 53 demonstrates the difference between the 8 and 7.5 min polishes. Eight minutes of polish were sufficient to clear all of the die, however, they are not planarized as well as the 7.5 min polish. In all cases a 2.5 min Phase II slurry polish was used to remove the barrier layer[3].

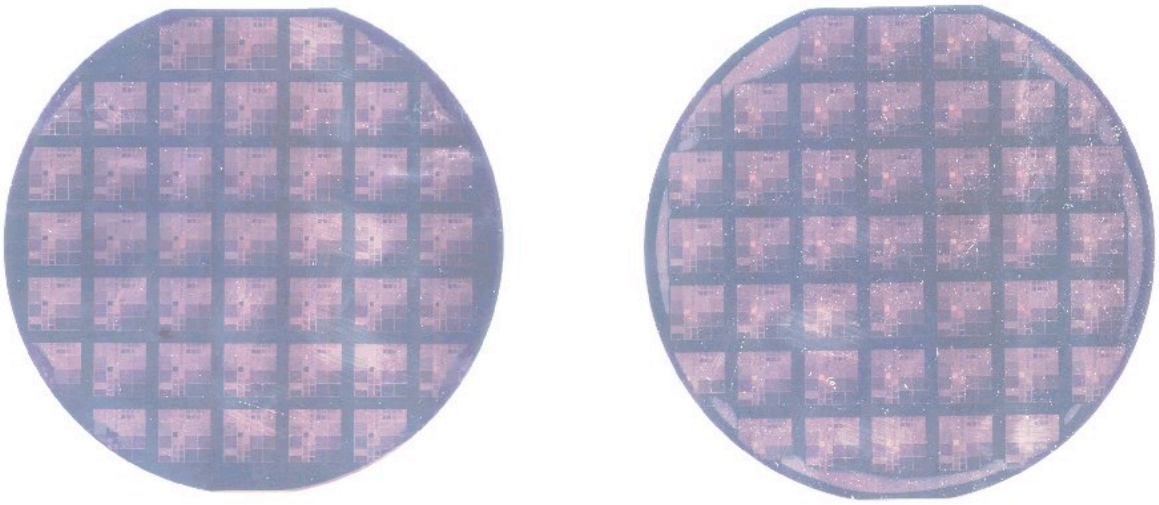


Figure 53: Remaining step heights of copper features after phase II slurry for 10% copper pattern density (45 µm oxide/5 µm Cu). 7.5 min polish has some copper remaining around the wafer edges.

[3] A 2.5 min polish with the second slurry was sufficient to remove the entire barrier layer.

Profilometer scans of 45/5 um feature from a polished wafer are shown in Figures 54 and 55. The degree of planarity is ((1 – 327/4700) x 100%) or 93%, which is sufficient for additional lithography pattering. This final topography is due to the removal of the barrier layer. The copper sticking up is an added benefit for the next layer of metal in terms of making good contact.

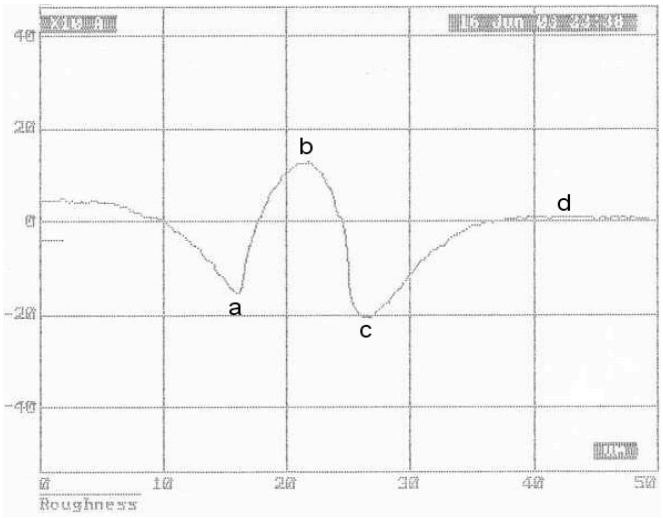


Figure 54: Profilometer results of oxide/copper features. Vertical distances are: ab = 275 Å, bc = 327 Å, bd = 116 Å (45 µm oxide/5 µm Cu feature)

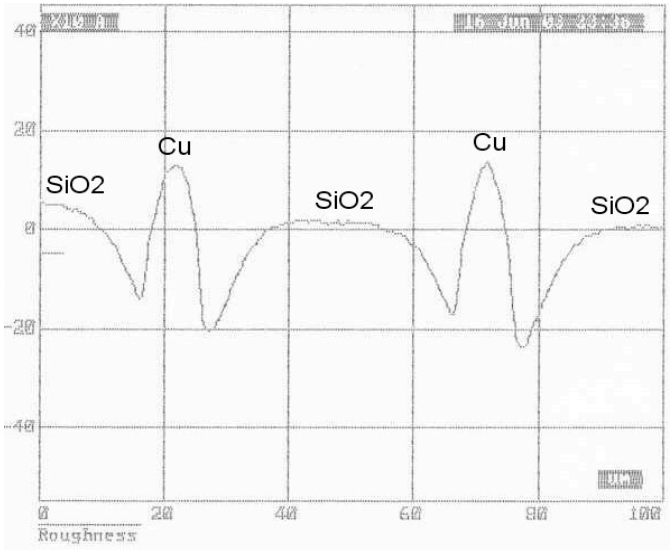


Figure 55: Profilometer results of oxide/copper features. Notice copper features are sticking above oxide level. (45 µm oxide/5 µm Cu feature).

Results from AFM scans are shown in Figures 56 and 57. Again topography appears to be on the order of 30 - 50nm. This demonstrates the ability of the CMP process to successfully planarized copper features. Furthermore, the process shows itself to be repeatable. The within die uniformity was ≤ 25 Å and the within wafer uniformity was ≤ 150 Å.

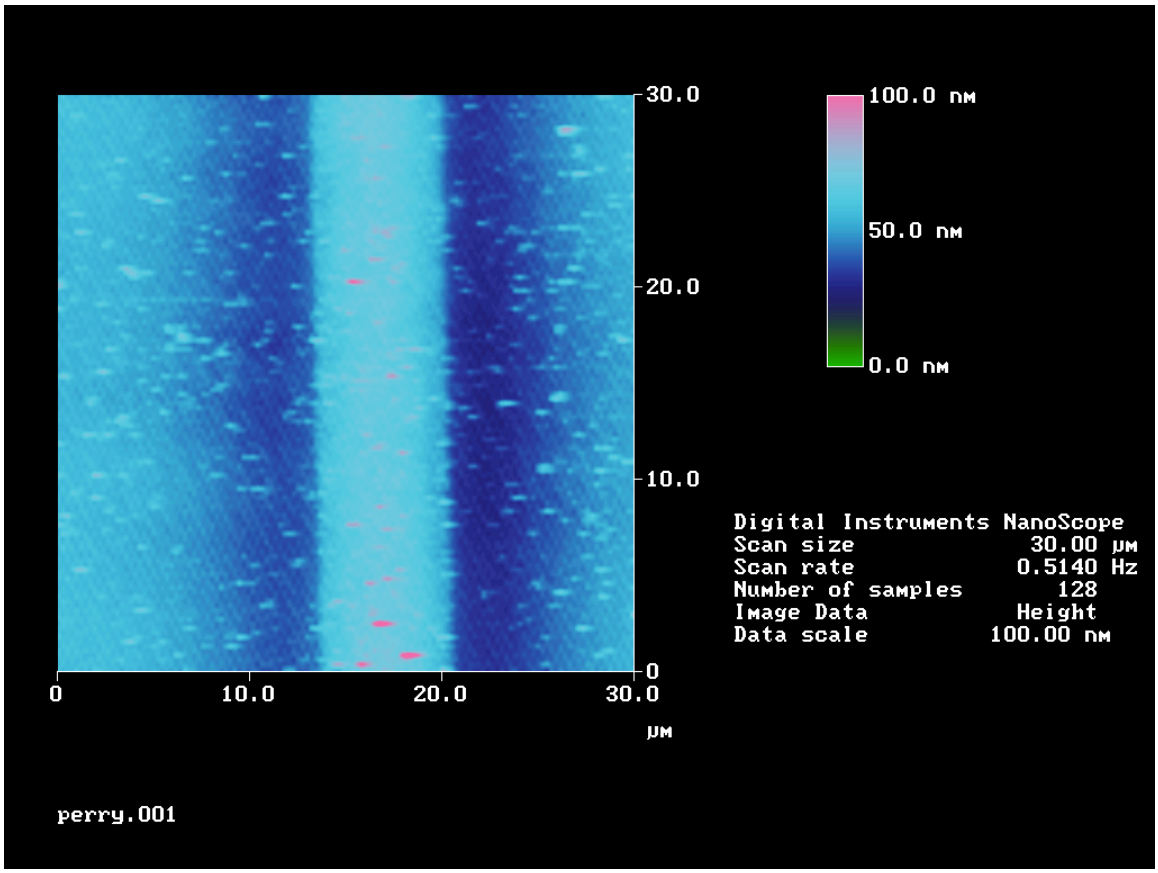


Figure 56: AFM image of oxide/copper features (45 µm oxide/5 µm Cu feature)

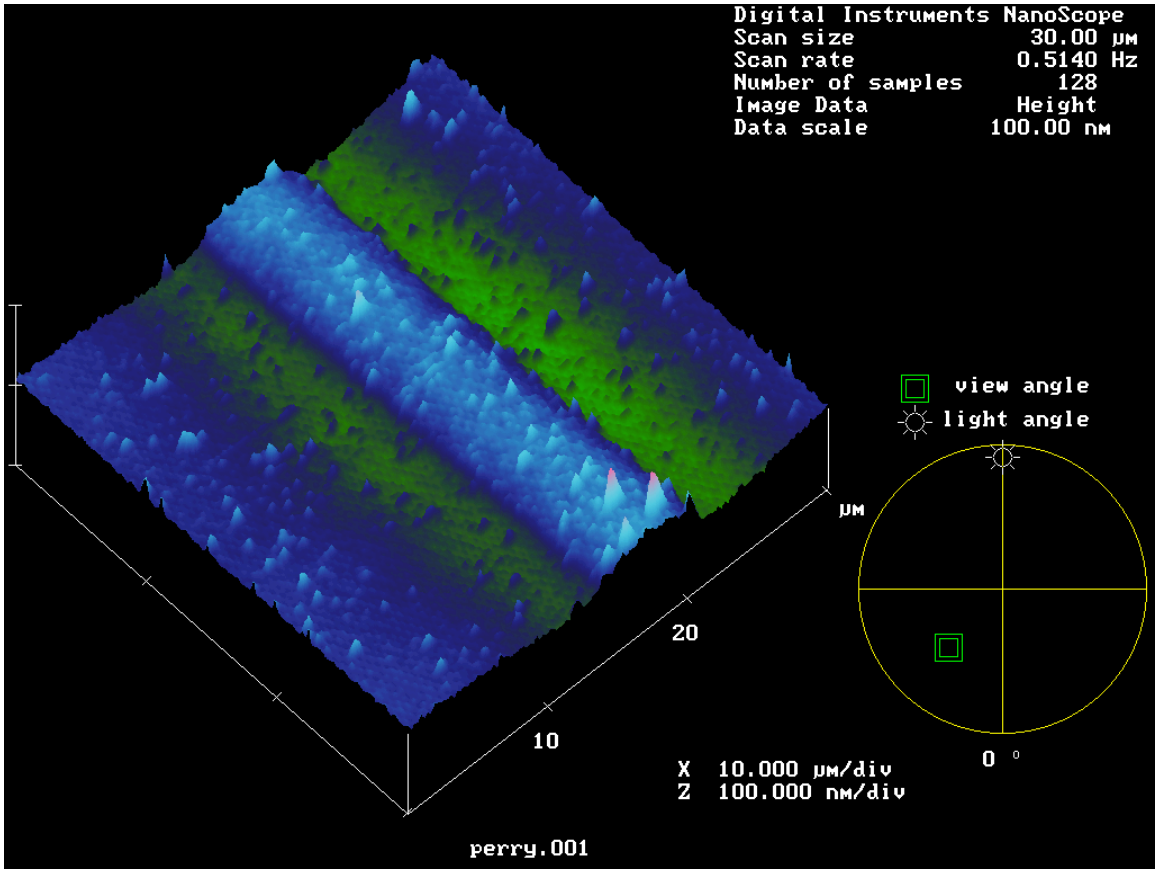


Figure 57: AFM image of oxide/copper features (45 µm oxide/5 µm Cu feature)

To summarize, the optimization of a copper damascene process will require that a characteristic CMP curve, as shown in Figure 52 needs to be constructed. This is necessary to limit the amount of copper dishing. In doing so, the following wafer parameters in Table 11 need to be fixed since altering them will have significant effects on the overall CMP performance.

Table 11: Some of the wafer parameters than can effect CMP optimization

Parameters
Feature size
Pattern density
Number die sites
Film stress
Feature step height
Barrier layer thickness [35]
Copper film thickness [36]

4.16 COPPER AND ALUMINUM MOS CAPACITORS

Device processing produced the following types of wafers/MOS capacitors:

a. 2 evaporated Al-MOS capacitors (Wafers 1 and 5)
b. 1 evaporated Al-MOS capacitors (with 250Å dry oxide on back, Wafer 9)
c. 2 sputtered Al-MOS capacitors (Wafers 3 and 7)
d. 2 Cu-MOS capacitors (Wafers Cu 1 and Cu 2)
e. 1 wafer with just a gate 250Å gate oxide

The data for the Cu and Al MOS capacitors tested to failure using the voltage ramp method are listed in Table 12. The details of this method are given in chapter 3 of reference [37]. Circular capacitor areas of $8x10^{-3}$ cm^2 were selected for testing so that the oxide capacitance was on the order of 1 nF. The SCA results for gate oxide characterization for the Al capacitors are given in Tables 13 and 14, where D_{it} is the interface trap density, Q_{fb} is the flat band charge and T_s is the minority carrier lifetime.

Table 12: Al and Cu MOS capacitor data

Maximum Electric Field (MV/cm)							
	Wafer 1	**Wafer 3**	**Wafer 5**	**Wafer 7**	**Wafer 9**	**Cu 1**	**Cu 2**
	Evap	**Sputter**	**Evap**	**Sputter**	**Evap w/ox**	**Sputter**	**Sputter**
% YIELD	64	52	12	12	4	32	68
AVG E_{max}	9.78	8.77	7.85	8.29	7.49	8.49	9.38
STDEV	2.25	1.59	1.44	1.30	0.75	2.21	1.33
% CV	23.0	18.2	18.3	15.7	9.9	26.0	14.1

Table 13: SCA data for Al capacitors after growing gate oxide

	Wafer 1	**Wafer 3**	**Wafer 5**	**Wafer 7**	**Wafer 9**
D_{it} ($cm^{-2}eV^{-1}$)	$8.89x10^{10}$	$8.18x10^{10}$	$8.48x10^{10}$	$8.91x10^{10}$	$8.76x10^{10}$
Q_{fb} (cm^{-2})	$1.08x10^{11}$	$1.22x10^{11}$	$1.02x10^{11}$	$1.01x10^{11}$	$1.03x10^{11}$
T_s (µsec)	111	124	129	124	130

Table 14: SCA data for wafer with only gate oxide grown on it

	Wafer after gate oxide	Wafer after being exposed to Al sintering step
D_{it} ($cm^{-2}eV^{-1}$)	$8.90x10^{10}$	$5.74x10^{10}$
Q_{fb} (cm^{-2})	$1.07x10^{11}$	$7.06x10^{10}$
T_s (µsec)	129	115

From the maximum electric field (E_{max}) results all die yields were poor with the exception of Wafers W1 and Cu 2 (above 60%). Al evaporated wafers, such as W1, should exhibit high yield rates, so W5, which faired no better than the Al sputtered ones, is an anomaly. Some plasma damage would be expected from the sputtered wafers and that should lower their yield [37]. Moreover, the Al evaporated wafer with gate oxide on the backside would be expected to have shown a higher E_{max} due to increased resistance of the back contact. Instead, it had the lowest average E_{max} and the worst yield. Our conclusion is that the back oxide should be removed to form a high quality ohmic contact.

The SCA data demonstrated reasonable levels for D_{it}, Q_{fb} and T_s after initial gate oxide growth which indicated that we had high quality oxides to start. After the blanket gate oxide wafer underwent the aluminum sintering, the interface traps and flat band charges were lowered. This is because the dangling bonds at the oxide interface were passivated. However, the minority carrier lifetime got worse. This was most likely due to metallic contaminates from the sinter tube and quartzware.

Much of the E_{max} data is skewed because of the particles of the wafer. This is evident in Figure 57 which shows the general trend of decreasing average E_{max} with increasing particle count. No statistical inferences due to processing variations could be drawn for the MOS capacitors. This was due to the large particle counts that skewed the data. The source of the high particle counts was traced to gas flow modifications that others had made to the oxide tube. We could not repeat this study until this contamination issue was resolved.

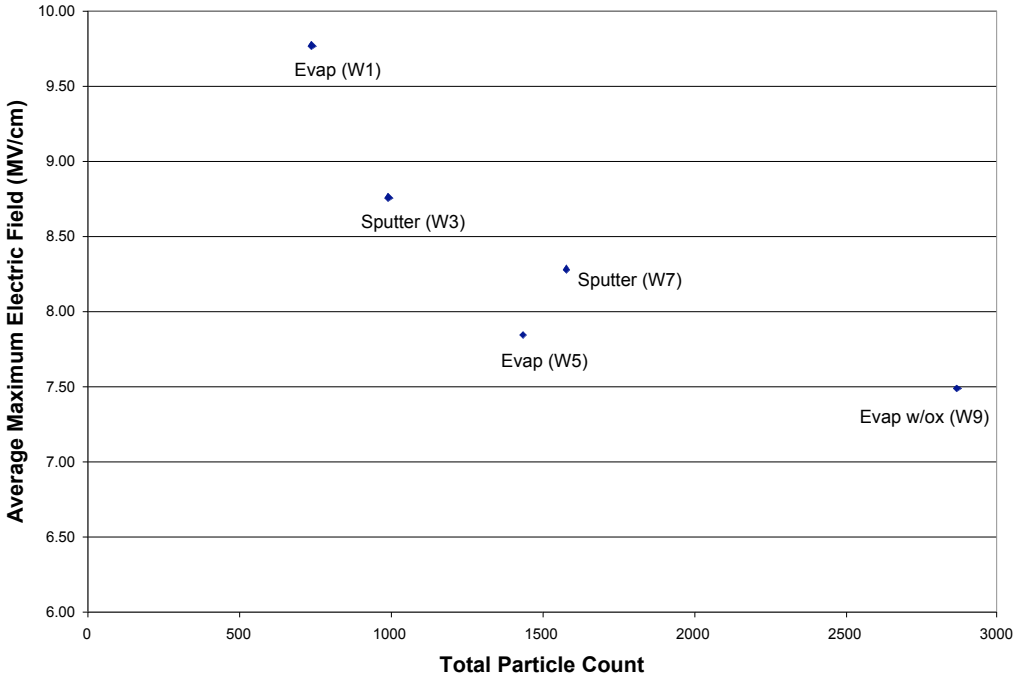


Figure 58: Average E_{max} vs. total particle count

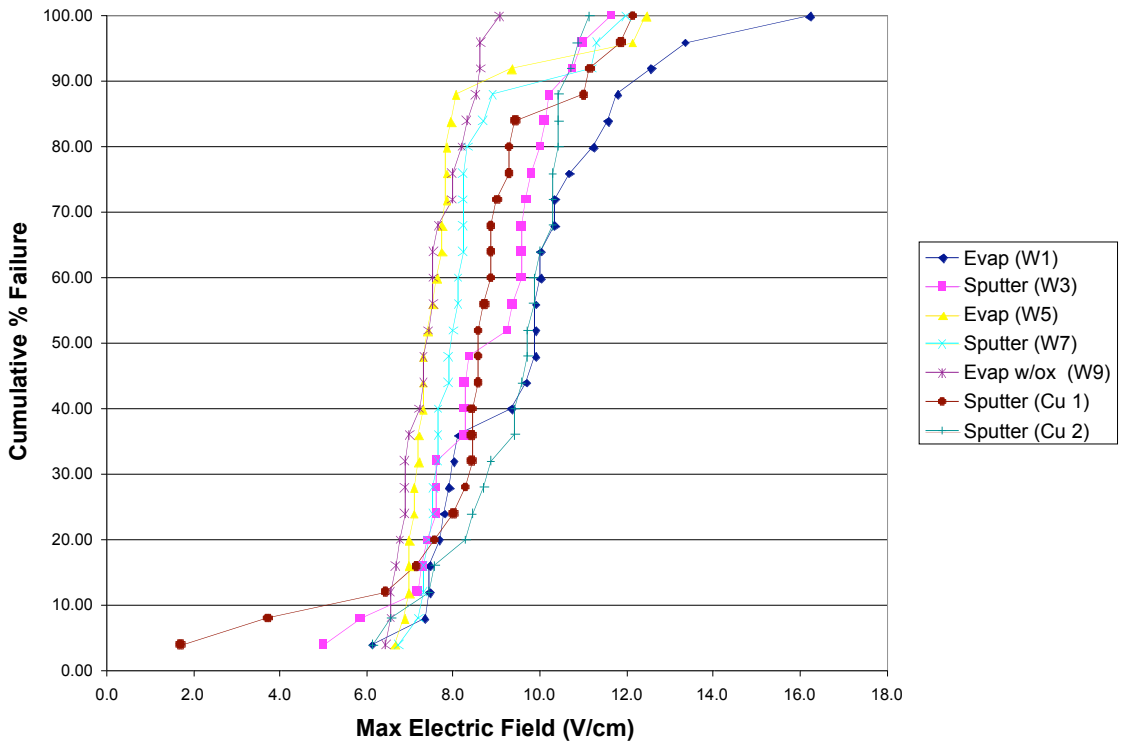


Figure 59: Cumulative failure plots for MOS capacitors shows the percent oxide failure from capacitor testing vs. maximum electric field

Cumulative failure plots in Figure 59 show that the Cu 1 and Cu 2 wafers processed via CMP were no better or worse than the conventionally processed wafers therefore; we concluded that CMP could be used for replacement gate processing. Furthermore, the wafer maps in Figure 60 fail to show any patterns in MOS capacitor failure, so again we believe the high particle counts are dominating the failure mechanisms and masking any effects due to processing.

Evap (W1)

8.00	7.33	6.11	7.89	13.33
8.11	7.78	7.67	7.44	7.44
10.33	9.33	12.56	9.89	10.00
10.67	9.89	10.00	9.67	9.89
16.22	11.22	11.56	10.33	11.78

Sputter (W3)

11.63	7.61	9.24	9.57	10.98
9.57	8.26	7.17	5.00	8.26
9.35	7.39	7.61	10.76	7.61
7.28	9.78	10.00	8.37	9.67
8.26	10.22	9.57	10.11	5.87

Evap (W5)

7.94	12.12	6.97	7.30	9.33
7.62	6.87	6.65	7.08	7.19
7.40	7.30	7.19	7.51	12.45
8.05	6.97	7.73	7.08	7.83
7.83	7.83	6.97	7.30	7.73

Sputter (W7)

11.19	7.65	7.31	7.19	11.99
7.31	7.53	6.74	7.99	8.68
7.65	8.22	8.33	8.22	7.42
11.30	7.88	7.88	7.65	8.22
8.22	8.11	7.53	8.90	8.11

Evap w/ox (W9)

6.55	6.44	7.31	8.62	8.30
7.64	7.97	8.19	7.53	6.99
7.53	6.55	7.42	6.88	7.53
7.31	6.77	7.97	6.88	7.21
9.06	8.52	8.62	6.66	6.88

Sputter (Cu 1)

7.57	8.43	8.86	3.71	6.43
8.00	8.57	8.57	9.00	8.29
8.86	8.43	9.43	9.29	9.29
7.14	11.14	8.86	1.71	8.57
8.71	11.00	12.14	11.86	8.43

Sputter (Cu 2)

10.29	10.00	9.71	9.86	10.43
10.29	10.29	6.57	11.14	9.43
8.43	7.43	9.86	7.57	10.86
10.43	10.71	9.57	8.71	8.29
10.43	8.86	9.43	9.71	6.14

Figure 60: Wafer maps for MOS capacitors. The values given are the maximum electric field values for each die. Dark areas indicate failure.

CHAPTER 5

CONCLUSIONS

Film stress in the Ta film was shown to be a function of sputtering pressure as well as film thickness. When deposited on SiO_2, tantalum formed the highly resistive tetragonal form and is in a state of compression. When deposited on top of TaN it exists in the same form but the film stack is less compressive in nature.

Target aging was shown to have a significant effect on Ta stress. Over time the discharge voltage drifted to lower values. The reduced voltage causes the film at low/moderate pressures to be less compressive. Conversely, the lowered voltage at high pressures causes Ta to be less tensile. A fairly large standard deviation in measured stress was typically observed when stress is determined from the change in substrate curvature before and after deposition. This is most likely due to film thickness non-uniformity.

A substantial increase in the CMP removal rate of Ta was shown to occur between approximately -400 and -1200 MPa of compressive stress. At these conditions the Ta has no benefit as a polish stop.

A bilayer of 100Å TaN/400Å Ta was successful in maintaining good adhesion to the copper. The reactively sputtered tantalum nitride for this bilayer was deposited with a 1:1 stoichiometry and existed in an amorphous state. This deposition process was conducted at 19.9 Watts/in^2, 12% N_2 (by flow rate) and 11.9 mTorr. Moreover, for different combinations of thickness for the 500 Å bilayer the stress was fairly constant at ≈ -690 MPa. It appears that bilayer stress is largely controlled by the TaN film.

The Ta deposition for this bilayer was carrier out at constant voltage. This was done to reduce the effects of target aging which can affect film adhesion properties due to fluctuating stresses in the film. Depositions were performed at 23.9 V/in^2 and 10 mTorr.

A repeatable copper CMP process was developed. For a given feature size (oxide/copper line-space pairs of 45/5 microns) it had a within die uniformity of ≤ 25 Å and a within wafer uniformity of ≤ 150 Å. It was determined that in order to optimize any given copper damascene process, a characteristic CMP curve of step height reduction vs. polish time needs to be constructed. This is necessary to limit the amount of copper dishing. To create a stable process wafer parameters such as feature size, pattern density, number of wafer die sites, film stress, feature step height, barrier layer thickness and copper film thickness need to be controlled.

Cu-MOS capacitors were fabricated to demonstrate the feasibility of replacement (damascene) metal gate technology at RIT. The capacitor yield data was greatly skewed by excessive particle counts. However, the Cu-MOS capacitors were comparable to the Al ones made under the same conditions, hence we do not observe any deleterious effects from the CMP process.

This work may be furthered by studying the following:

- Copper stress vs. depositions conditions and/or electroplated copper
- Dual damascene copper process
- Repeating the capacitor study

REFERENCES

1. Pai P.L. and Ting C.H., *Selective electroless copper for VLSI interconnection***,** IEEE Electron Device Letters, 10(9): 423-425 (1989).

2. Jeng S.P., Havemann R.H., and Chang M.C., *Process integration and manufacturability issues for high performance multilevel interconnect*, Materials Research Society Symp. Proc. Vol. 337, 25-31 (1994).

3. Plummer J.D., Deal M.D., Griffin P.B., *Silicon VLSI Technology Fundamentals, Practice and Modeling*, Prentice Hall, Upper Saddle River, NJ, 2000.

4. Kaloyeros, A.E. and Eisenbraun, E.*, Ultra diffusion barriers/liners for gigascale copper metallization*, Annu. Rev. Mater. Sci. 30: 363-385 (2000).

5. Holloway, K. et al., *Tantalum as a diffusion barrier between copper and silicon. Failure mechanism and effect of nitrogen additions*, Journal of Applied Physics, 71(11): 5433-5443 (1992).

6. CallisterW.D., *Fundamentals of Materials Science and Engineering*, John Wiley & Sons, New York, 2001.

7. Pierson H.O., *Handbook of Refractory Carbides and Nitrides*, Noyes Publications, Westwood, NJ, 1996.

8. Ryu, C., Lee, H., Kwon, K., Loke, A., and Wong, S.*, Barriers for copper interconnections*, Solid State Technology, April 1999, p. 53-56.

9. Peters, Laura, *Finding the ultimate copper barrier and seed*, Semiconductor International, July 2001, p. 23

10. Chin, B.L. et al., *Barrier and seed technologies for sub-0.10 μm copper chips*, Semiconductor International, May 2001, p. 107-114.

11. Lane M., Dauskardt H.*, Adhesion and reliability of copper interconnects with Ta and TaN barrier layers*, Journal of Materials Research, 15(1): 203-211 (2000).

12. Milton Ohring, *The Materials Science of Thin Films*, 2nd edition, Academic Press, New York, 2002.

13. Aksu, Serdar, *The role of complexing agents in the chemical mechanical planarization (CMP) of copper thin films*, Ph.D. Thesis Proposal- University of California at Berkeley (2000).

14. Lou Q., Ramarajan, S., Babu and S.V., *Modification of the Preston equation for the chemical-mechanical polishing of copper*, Thin Solid Films, 335: 160-167 (1998).

15. Wrschka P., Hernandez J., Oehrlein G.S. and King, J., *Chemical mechanical planarization of copper damascene structures*, Journal of The Electrochemical Society, 147(2): 706-712 (2000).

16. Thornton, J.A., Hoffman, D.W., *Stress related effects in thin films*, Thin Solid Films, 171: 5-31 (1989).

17. Windischmann, H., *Intrinsic stress in sputtered-deposited thin films*, Critical Reviews in Solid State and Material Science, 17(6): 547-596 (1992).

18. Targove, J.D. and Macleod, H.A., *Verification of momentum transfer as the dominant densifying mechanism in ion-assisted deposition*, Appl. Opt., 27, 3779 (1988).

19. Hoffman D.W., *Perspective on stresses in magnetron-sputtered thin films*, J. Vac. Sci. Technol. A 12(4), 953-961 (1994).

20. Pan, J., Ngo, M., Woo, C., Goo J., Besser P., Yu B., Xiang Q., Lin M., *Metal gate NMOSFETS with TaSiN/TaN stacked electrode fabricated by a replacement (damascene) technique*, Technology Research Group, Advanced Micro Devices (2001).

21. Gould S.A.C., Drake C.B., Prater A.L. et al., The atomic force microscope: a tool for science and industry, Ultramicroscopy, 33: 93-98 (1990).

22. Omar M.A., *Elemantary Soid State Physics*, Addison Wesley Longman, New York, 1993.

23. Kern, W. Editor, *Handbook of semiconductor wafer cleaning technology*, Noyes Publications, Park Ridge, New Jersey, 497-516, 1993.

24. Thornton, J.A., Hoffman, D.W., *The influence of discharge current on the intrinsic stress in Mo films deposited using cylindrical and planar magnetron sputtering sources*, J. Vac. Sci. Technol. A 3(3), 576-579 (1985).

25. Tseng, W., Wang, Y., Niu, J., *Microstructure-related resistivity change after chemical-mechanical polish of Al and W thin films*, Thin Solid Films, 370, 96-100 (2000).

26. Tseng, W. and Wang, Y., *The intercorrelation between microstructure and chemical-mechanical polish of metal thin films*, Mat. Res. Soc, Proc. 564, 459-464, (1999).

27. Jindal, A., Li, Y., Babu, S.V., *Effect of pH on chemical-mechanical polishing of copper and tantalum*, Mat. Res. Soc, Proc. 671, M6.8.1 – M6.8.6, (2001).

28. Ramarajan, S., Li, Y., Hariharaputhiran, M., Her, Y.S. and Babu, S.V., *Effect of pH and ionic strength on chemical mechanical polishing of tantalum*, Electrochemical and Solid-State letters, 3(5), 232-234, (2000).

29. Riekkinen T., Molarius J. et al., *Reactive sputter deposition and properties of Ta_xN thin films*, Microelectronic Engineering, 64, 289-297, (2002).

30. Nie H.B., Xu S.Y. et al., *Structural and electrical properties of tantalum nitride thin films fabricated by using reactive radio-frequency magnetron sputtering*, Applied Physics A: Material Science & Processing, 73, 229-236, (2001).

31. Sun, X., Kolawa E., et al., *Properties of reactively sputter-deposited Ta-N thin films*, Thin Solid Films, 236, 347-351, (1993).

32. Lin J., Chen, G. et al., *Growth of tantalum nitride films on Si by radio frequency reactive sputtering: effect of N_2/Ar flow ratio*, Journal of The Electrochemical Society, 146(5), 1835-1839, (1999).

33. Maissel L., Glang R., *Handbook of Thin Film Technology*, McGraw-Hill, New York, 1970.

34. Stavreva Z., Zeidler D. et al., *Chemical-mechanical polishing of copper for interconnect formation*, Microelectronic Engineering, 33, 249-257, (1997).

35. Hansen D., Moloney G., Witty, M., *Copper CMP: The role of barrier material and its effect on dishing and oxide erosion*, Electrochemical Society Proceedings, 99, 136-148 (1999).

36. Tsai T.C., Chen H.C. et al., *Effects of copper film thickness on copper CMP performance*, Electrochemical Society Proceedings, 26, 269-272 (2000).

37. Capasso, Keith, *Process development and reliability of thin gate oxides*, M.S. Thesis - Rochester Institute of Technology (1999).

38. Powell R.A. & Rossnagel S.M., *PVD for microelectronics: sputter deposition applied to semiconductor manufacturing*, Academic Press, New York, 1999.

39. Hari Singh Nalwa, *Deposition and processing of thin films*, Vol 1, Academic Press, New York, 2002.

APPENDIX

1. XRD report for tantalum and tantalum nitride films. Attached is an original copy of the report. Therefore, the figures are without titles/captions. This was done to preserve in its original form of the report. (p. 78-83)

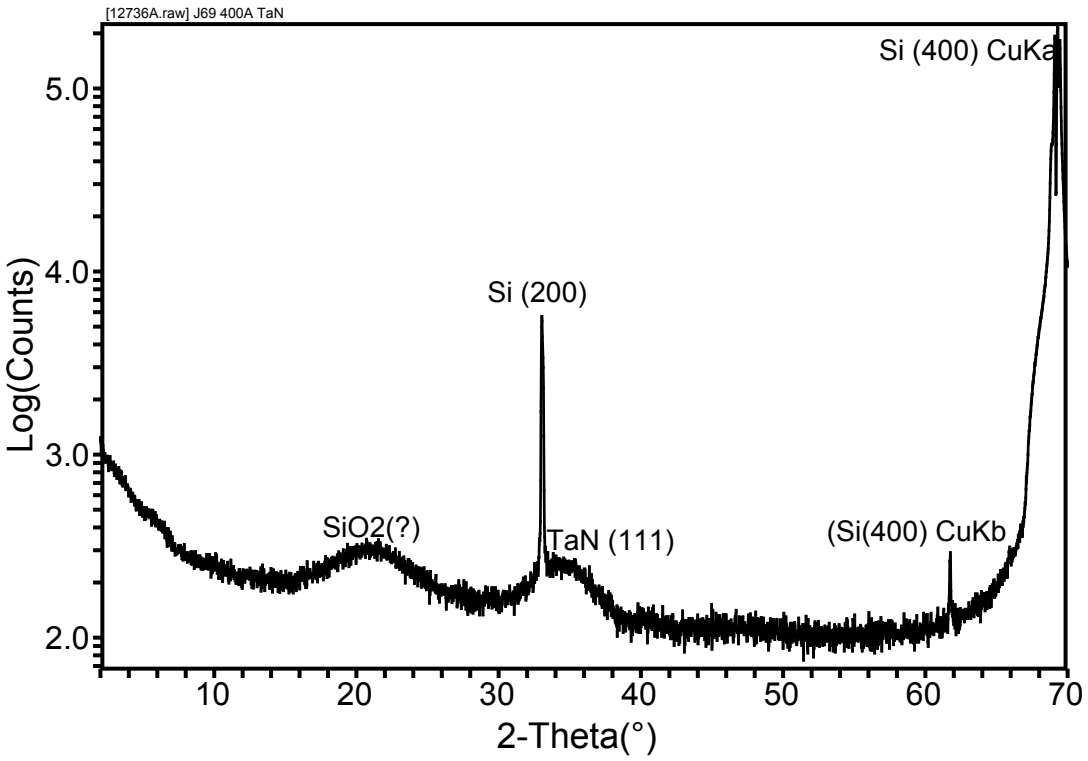
2. Process steps for fabricating aluminum MOS capacitors. (p. 84)

To: Chris Hoople/Jeffrey Perry
From: Tom Blanton
Subject: XRD analysis of Ta/TaN thin film samples
Job 12736 June 24, 2003

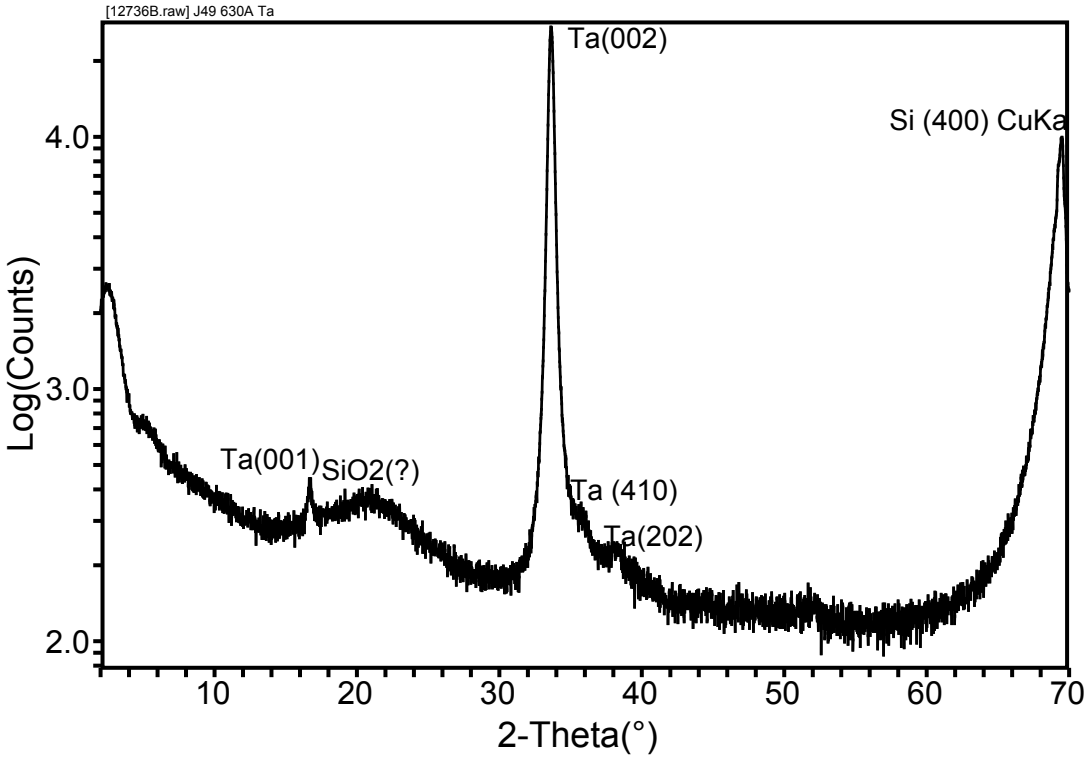
Three samples of thin films deposited on (100) Si were submitted to XRD for characterization. Samples were analyzed using a Rigaku D2000 diffractometer equipped with a copper rotating anode, diffracted beam graphite monochromator tuned to CuKα radiation, and a scintillation detector. Note that the diffraction patterns are plotted on a log intensity scale due to the very intense Si(400) diffraction peak. Samples aligned with near perfect orientation will show the Si(200) peak.

Suggested references: Stavrev et al., Thin Solid Films, 307, pp79-88. (1997) and Thesis from Deepa Gazula, Rochester Institute of Technology

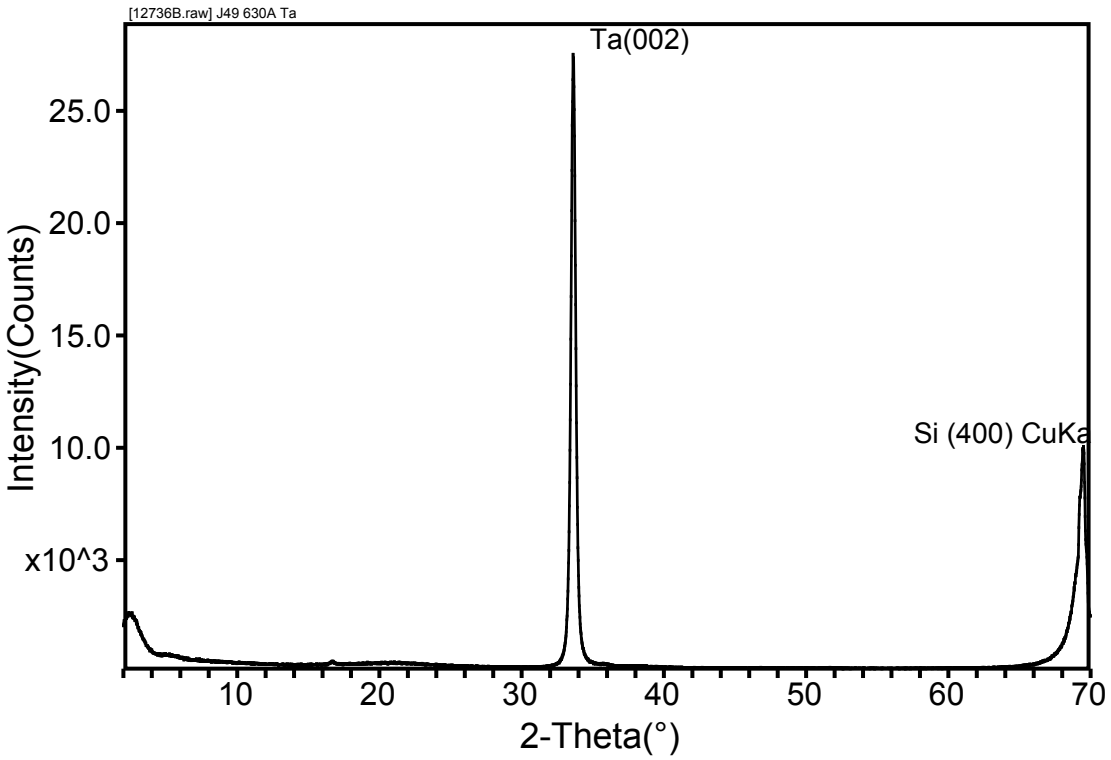
Sample J69, 400 Å TaN: There is an amorphous peak at ~34 ° 2θ which is consistent with the (111) TaN peak. There is also an amorphous peak at ~ 21 ° 2θ which may be due to amorphous SiO2. There is no Ta_xN_y peak expected at this 2θ position.



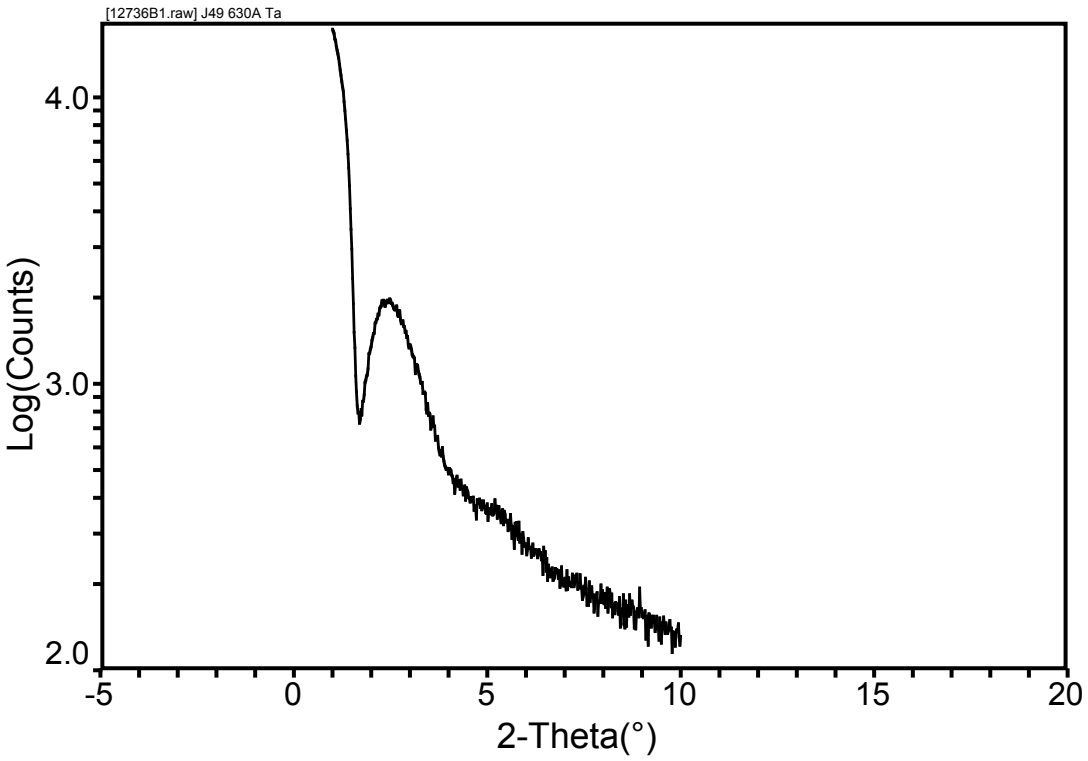
Sample J49, 630 Å Ta:
This sample shows the presence of highly oriented β - Ta, showing (00L) planar orientation (β - Ta is tetragonal, PDF Number 25-1280). The Ta (001) peak is likely a Ta (002) _/2 harmonic peak. There is also an amorphous peak at ~ 21 ° 2θ which may be due to amorphous SiO2. Using the Scherrer technique, the nominal (002) β - Ta crystallite size is 230 Å.



For illustration purposes a linear intensity scale plot is shown below for J49, demonstrating the well aligned (00L) β - Ta.

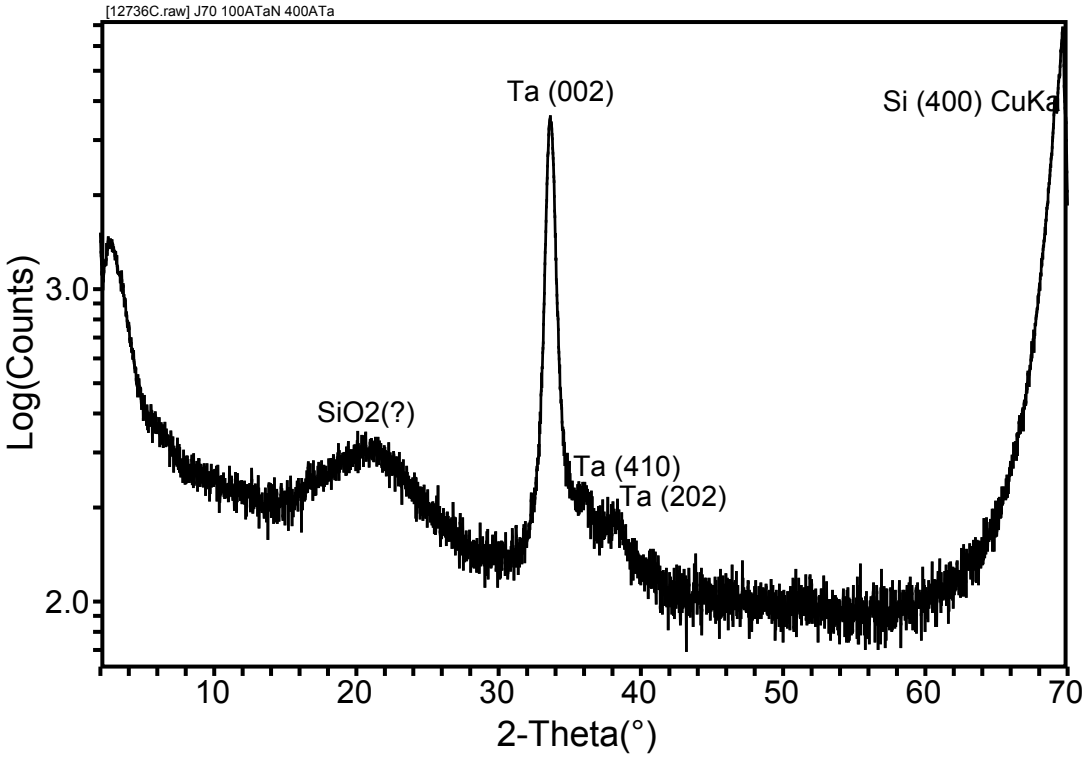


Also note that there appears to be a low angle peak at ~2° 2θ. The low angle plot below (using tighter slits) shows that there is some indication of an interference fringe due to film thickness. There are not enough fringes observed to do a thickness calculation but the presence of the fringes does indicate that the film thickness is reasonably uniform.

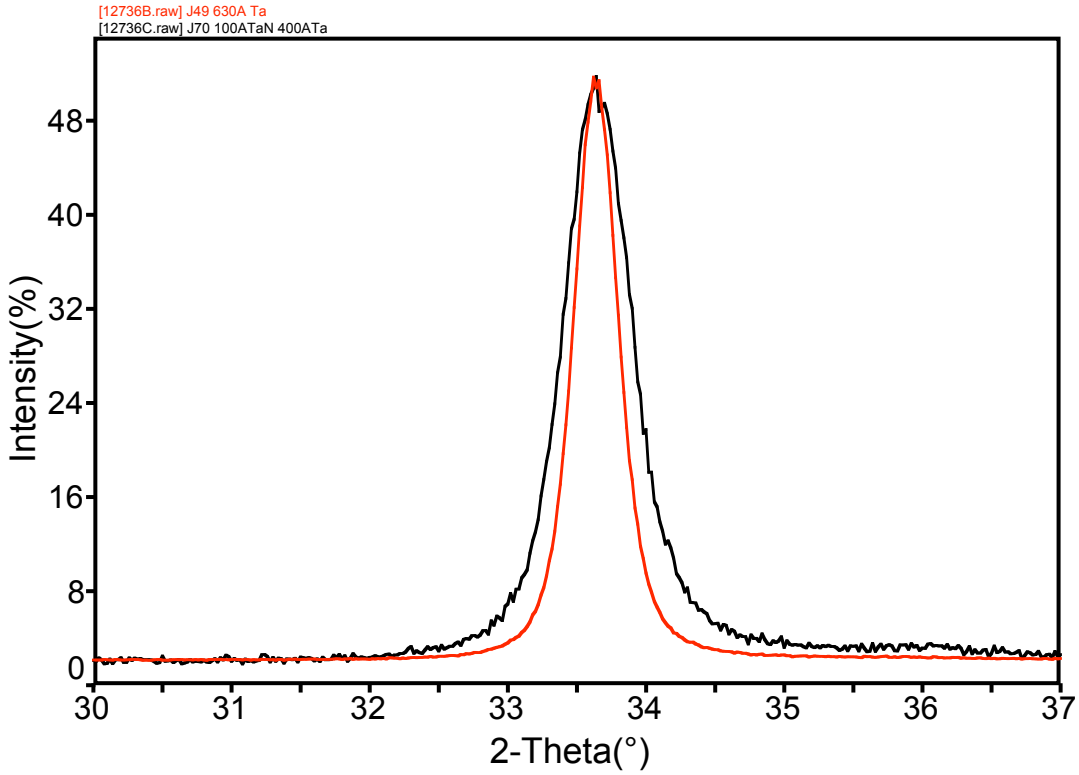


Sample J70, 100 Å TaN/400 Å Ta:

This sample shows the presence of highly oriented β - Ta, showing (00L) planar orientation (β - Ta is tetragonal, PDF Number 25-1280). Identification of TaN (assumed to be amorphous like J69) is hindered by the presence of the β - Ta peaks. There is also an amorphous peak at ~ 21 ° 2θ which may be due to amorphous SiO2. A low angle peak at ~2 ° 2θ is consistent with a reflectivity fringe as seen in J49 above. Using the Scherrer technique, the nominal (002) β - Ta crystallite size is 160 Å.



Overlapping the two Ta (002) peaks after normalizing intensity shows J49 has a narrower diffraction peak than J70, due to larger crystallite size.



In summary, the TaN films are amorphous, the Ta films are crystalline and show (00L) planar orientation.

Process Steps for Making MOS Capacitors

1. Select 8 device wafers and 4 dummy to grow 5000Å of wet oxide
2. Determine particle count before RCA clean
3. RCA clean
4. Determine particle count after RCA clean
5. Grow 5000Å of field oxide
6. Determine particle count after growing field oxide
7. Measure oxide thickness
8. Run 9-pt SCA pattern on all 8 wafers. How similar are the substrates?
9. Perform lithography on 2 best wafers (darkfield capacitor mask)
10. Etch all wafers in clean BOE
11. Strip photoresist off patterned wafers using acetone
12. Determine particle count
13. RCA clean all wafers
14. Determine particle count
15. Preclean tube with chlorine preclean process. Grow 250Å gate oxide.
16. Determine particle count
17. Measure oxide thickness
18. SCA the 6 blanket gate oxide wafers
19. If SCA results are satisfactory deposit TaN/Ta/Cu on patterned wafers using standard sputtering recipe.
20. Evaporate Al onto 3 of blanket oxide wafers. Use two Al pellets. This should deposit approximately 0.75 μm of Al.
21. Sputter Al onto 2 of blanket oxide wafers. Run at 1500 W for 15 min. This will deposit 0.55 μm of Al.
22. Perform lithography on all 5 Al coated wafers (lightfield capacitor mask) and etch Al features
23. Polish patterned copper wafers to make capacitors
24. Coat resist on all front side of patterned wafers and bake at 90^{o}C for 1 min using Shipley 812 resist
25. Etch backside of 6 patterned wafers except one of the evaporated Al wafers. Use BOE for 30 sec etch.
26. Sputter Al on the backside of all 7 patterned wafers. Don't utilize a bake-out step. Run at 1500 W for 15 min to get 0.55 μm of Al.
27. Sinter all Al patterned wafers. Do for 20 min at 400^{o}C with H_2/N_2 forming gas.
28. Sinter copper wafers in Heraeus vacuum oven at 350^{o}C for 60 min. Perform two nitrogen purges at room temperature and a third at 125^{o}C. This should be enough to remove any residual water vapor.
29. In the end should have 8 wafers: 5 are for Al-MOS Caps, 2 are Cu-MOS Caps, and 1 is just the gate oxide

    f. 2 evaporated Al-MOS Caps
    g. 1 evaporated Al-MOS Cap (with oxide on back)
    h. 2 sputtered Al-MOS Caps
    i. 2 Cu-MOS Caps
    j. 1 wafer with just gate oxide (with oxide on back)

Note: For both RCA cleans and oxide growths insure that the dummy wafers "box-in" the device wafers. The dummy wafers should have the same spacing as the device wafers. D1-D2-W1-W2-W3-W4-W5-W6-W7-W8-D3-D4